Statistics

INVARIANT TESTS AND LIKELIHOOD RATIO TESTS FOR MULTIVARIATE ELLIPTICALLY CONTOURED DISTRIBUTIONS

TECHNICAL REPORT NO. 14

HUANG HSU

MAY 1985

U. S. ARMY RESEARCH OFFICE
CONTRACT DAAG29-82-K-0156
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

INVARIANT TESTS AND LIKELIHOOD RATIO TESTS FOR MULTIVARIATE ELLIPTICALLY CONTOURED DISTRIBUTIONS

Technical Report No. 14

Huang Hsu
Stanford University

May 1985

U. S. Army Research Office
Contract DAAG29-82-K-0156

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my advisor Professor Theodore W. Anderson for his constant guidance, continual support and encouragement that have made this report possible during the period of my studies at Stanford. I would also like to thank Professors Charles Stein and Ingram Olkin for reading this manuscript and making many helpful suggestions. The "Chinese Mafia" have been important to me. Thanks are due to Drs. Chen-Hsin Chen, Inchi Hu, Young-Lim Su, Bwo-Han Yung, Sweetsun Chen, Chung-Teh Wu, Lian-Tai Chen, and Gary Chang. I am grateful to my parents and my fiancee, Men-Li, for their moral support and care they have given me all through the years. I dedicate this report to them.

Contents

Abstract

The usual assumption in multivariate hypothesis testing is that the sample consists of n independent, identically distributed Gaussian p-vectors. In this dissertation this assumption is weakened by considering a class of distributions for which the vector observation are not necessarily Gaussian or independent. This class consists of the elliptically symmetric laws with densities of the form $f(\boldsymbol{X}_{n\times p}) = g(\mathrm{tr}(\boldsymbol{X}-\boldsymbol{M})'(\boldsymbol{X}-\boldsymbol{M})\boldsymbol{\Sigma}^{-1})$. The following hypothesis testing problems are considered: testing for equality between the mean vector and a specified vector, lack of correlations among different sets, equality of covariance matrices and mean vectors, equality between the correlation coefficient and a specified number, and MANOVA. For each of the above hypotheses, invariant tests and their properties are developed. These include the uniformly most powerful test, the locally most powerful test, admissibility, and null and non-null distributions. Further, under the assumptions that $g(\cdot)$ is continuous, each element of the covariance matrix of $\boldsymbol{X}$ is finite, and the null hypothesis is scalar-invariant, it is shown that the usual normal-theory likelihood ratio tests are exactly robust for the null case under this wider class (i.e. the likelihood ratio tests, sampling from this general class, are the same as the usual normal-theory likelihood ratio tests and their null distributions are the same.)

Chapter 1

INTRODUCTION

The usual assumption in multivariate hypothesis testing is that the sample consists of independent and identically distributed Gaussian vectors. In this paper we consider the more general hypothesis testing problem when the sample observations, not necessarily either Gaussian or independent, are from the family of multivariate elliptically contoured distributions.

The elliptically contoured distributions on the n-dimensional Euclidean space $\Re^n$ are defined as follows. If the characteristic function of an n-dimensional random vector $\mathbf{x}$ has the form $\exp(i\mathbf{t}'\mu)\phi(\mathbf{t}'\Sigma^*\mathbf{t})$, where $\mu : n \times 1$, $\Sigma^* : n \times n$, $\text{rank}(\Sigma^*) = k$, $\Sigma^* \geq 0$, and $\phi \in \Phi_k = \{ \phi \mid \phi(\cdot) \text{ is a function such that } \phi(t_1^2 + \cdots + t_k^2) \text{ is a characteristic function on } \Re^k \}$, we say that $\mathbf{x}$ is distributed according to an elliptically contoured distribution with parameters μ, Σ^*, and ϕ, and write $\mathbf{x} \sim EC_n(\mu, \Sigma^*, \phi)$.

Elliptically contoured distributions have been extended to the case of matrices by Dawid (1977, 1978), Chmielewski (1980), and Anderson and Fang (1982b).

Let $\mathbf{X}$, $\mathbf{M}$, and $\mathbf{T}$ be $n \times p$ matrices. We express them in terms of elements, columns, and rows as

$$\mathbf{X} = (x_{ij}) = (\mathbf{x}_1, \ldots, \mathbf{x}_p) = \begin{pmatrix} \mathbf{x}'_{(1)} \\ \vdots \\ \mathbf{x}'_{(n)} \end{pmatrix}, \quad \mathbf{x} = \operatorname{vec} \mathbf{X}',$$

$$\mathbf{M} = (\mu_{ij}) = (\mu_1, \ldots, \mu_p) = \begin{pmatrix} \mu'_{(1)} \\ \vdots \\ \mu'_{(n)} \end{pmatrix}, \quad \mu = \operatorname{vec} \mathbf{M}', \tag{1.1}$$

$$\mathbf{T} = (t_{ij}) = (\mathbf{t}_1, \ldots, \mathbf{t}_p) = \begin{pmatrix} \mathbf{t}'_{(1)} \\ \vdots \\ \mathbf{t}'_{(n)} \end{pmatrix}, \quad \mathbf{t} = \operatorname{vec} \mathbf{T}',$$

where $\mathbf{x} = \operatorname{vec} \mathbf{X}' \stackrel{def}{=} \left(\mathbf{x}'_{(1)}, \ldots, \mathbf{x}'_{(n)}\right)'$ with the corresponding meanings for μ and $\mathbf{t}$.

If the characteristic function of a random matrix $\mathbf{X}$ has the form

$$\exp\left(i \sum_{j=1}^{n} \mathbf{t}'_{(j)} \mu_{(j)}\right) \phi\left(\mathbf{t}'_{(1)} \Sigma_1 \mathbf{t}_{(1)}, \ldots, \mathbf{t}'_{(n)} \Sigma_n \mathbf{t}_{(n)}\right), \tag{1.2}$$

with $\Sigma_1, \ldots, \Sigma_n \geq 0$, we say that $\mathbf{X}$ is distributed according to a multivariate (rows) elliptically contoured distribution (MECD) and write $\mathbf{X} \sim MEC_{n \times p}(\mathbf{M}; \Sigma_1, \ldots, \Sigma_n; \phi)$. In this paper we consider only the subclass of MECD in which the function ϕ satisfies

$$\phi(t_1, \ldots, t_n) = \varphi(t_1 + \cdots + t_n). \tag{1.3}$$

We continue to denote MECD in this subclass by $MEC_{n \times p}(\mathbf{M}; \Sigma_1, \ldots, \Sigma_n; \phi)$ and φ by ϕ.

When $\mu_{(1)} = \mu_{(2)} = \cdots = \mu_{(n)} = \mu$ and $\Sigma_1 = \cdots = \Sigma_n = \Sigma$, we write $\mathbf{X} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$, and if the density of $\mathbf{X}$ exists (in this paper, this is the usual assumption), then it has the form

$$|\Sigma|^{-\frac{1}{2}n} g\left(\operatorname{tr} \Sigma^{-1} \mathbf{G}\right), \tag{1.4}$$

where $\mathbf{G} = \sum_{j=1}^{n} \left(\mathbf{x}_{(j)} - \mu\right)\left(\mathbf{x}_{(j)} - \mu\right)'$.

Let $\mathbf{u}^{(q)}$ denote a random vector which is uniformly distributed on the unit sphere in $\Re^q$ and $\Omega_q\left(\|\mathbf{t}\|^2\right)$ denote its characteristic function. Schoenberg (1938) pointed out that a characteristic function $\phi \in \Phi_q$ if and only if

$$\phi(t) = \int_0^\infty \Omega_q\left(tr^2\right)\, dF(r) \tag{1.5}$$

for some distribution function F on $[0, \infty)$.

Throughout $\mathbf{X} \stackrel{d}{=} \mathbf{Y}$ means that the random matrices $\mathbf{X}$ and $\mathbf{Y}$ are identically distributed; rank($\mathbf{A}$) denotes the rank of the matrix $\mathbf{A}$.

Cambanis, Huang and Simons (1981) obtained the following two properties :

(i) $\mathbf{x} \sim EC_n(\mu, \Sigma^*, \phi)$ with $\text{rank}(\Sigma^*) = k$ if and only if

$$\mathbf{x} \stackrel{d}{=} \mu + R\mathbf{D}'\mathbf{u}^{(k)}, \tag{1.6}$$

where $R \geq 0$ is independent of $\mathbf{u}^{(k)}$, $\Sigma^* = \mathbf{D}'\mathbf{D}$ is a factorization of Σ^* (i.e., $\mathbf{D}$ is a $k \times n$ matrix and $\text{rank}(\mathbf{D}) = k$) and the distribution function F of R is related to ϕ as in (1.5) with k substituted for q. For convenience we denote this relationship by $R \leftrightarrow \phi \in \Phi_k$. Here $\mathbf{x} \stackrel{d}{=} \mathbf{y}$ denotes that the random vectors $\mathbf{x}$ and $\mathbf{y}$ are identically distributed.

(ii) Write $\mathbf{u}^{(n)} = \left(\mathbf{u}_1^{(n)\prime}, \mathbf{u}_2^{(n)\prime}\right)'$, where $\mathbf{u}_1^{(n)}$ is m-dimensional column vector $(1 \leq m < n)$. Then $\left(\mathbf{u}_1^{(n)}, \mathbf{u}_2^{(n)}\right) \stackrel{d}{=} \left(R_{mn}\mathbf{u}^{(m)}, (1 - R_{mn}^2)^{\frac{1}{2}}\mathbf{u}^{(n-m)}\right)$, where $R_{mn} (\geq 0)$, $\mathbf{u}^{(m)}$, and $\mathbf{u}^{(n-m)}$ are independent, and $R_{mn}^2 \sim \text{Beta}\left(\frac{m}{2}, \frac{n-m}{2}\right)$.

Anderson and Fang (1982b) pointed out that $\mathbf{X} \sim MEC_{n\times p}(\mathbf{M}; \Sigma_1, \ldots, \Sigma_n; \phi)$ if and only if $\mathbf{x} = \text{vec}\,\mathbf{X}' \sim EC_{np}(\mu, \mathbf{V}, \phi)$ where $\mathbf{V}$ is a matrix with diagonal blocks Σ_i, $i = 1, \ldots, n$, and off-diagonal blocks $\mathbf{0}$, and $\mu = \text{vec}\,\mathbf{M}'$. So it is easy to see that if $\mathbf{X} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$ and $\text{rank}(\Sigma) = \ell$ then

$$\mathbf{X} \stackrel{d}{=} \mathbf{1}_n\mu' + R\mathbf{U}\mathbf{D}, \tag{1.7}$$

where $\mathbf{1}_n' = (1, 1, \cdots, 1)$, $\mathbf{U} : n \times \ell$, $\text{vec}\,\mathbf{U} = \mathbf{u}^{(n\ell)}$, $\mathbf{D} : \ell \times p$, $\mathbf{D}'\mathbf{D} = \Sigma$, and $R \leftrightarrow \phi \in \Phi_{n\ell}$.

When $\Sigma > 0$, $\mathbf{X} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$ with $\phi(t) = \exp(-\frac{1}{2}t)$ if and only if $\mathbf{x}_{(j)}$, $j = 1, \ldots, n$, are i.i.d. $N_p(\mu, \Sigma)$. (This implies that sampling from a multivariate normal distribution is a special case of sampling from the family of multivariate elliptically contoured distributions.)

There are three definitions to be emphasized here :

(i) The Dirichlet Distribution $D_m(p_1, \cdots, p_{m-1}; p_m)$.

If $\mathbf{y} = (y_1, \cdots, y_m)'$ is a random vector with $\sum_{i=1}^m y_i = 1$ and $(y_1, \cdots, y_{m-1})'$ has the density with $p_i > 0$, $i = 1, \ldots, m$,

$$f_m(t_1, \cdots, t_{m-1}) = \begin{cases} \frac{\Gamma(\sum_{i=1}^m p_i)}{\prod_{i=1}^m \Gamma(p_i)} \prod_{i=1}^{m-1} t_i^{p_i-1} \left(1 - \sum_{i=1}^{m-1} t_i\right)^{p_m-1}, & \text{if } t_i \geq 0, \\ & i = 1, \cdots, m-1, \\ & \sum_{i=1}^{m-1} t_i \leq 1; \\ 0, & \text{otherwise,} \end{cases} \tag{1.8}$$

we say that $(y_1, \cdots, y_{m-1})' \sim D_m(p_1, \cdots, p_{m-1}; p_m)$.

(ii) The distribution $MG_{p,k+1}(\Sigma; \frac{n_1}{2}, \cdots, \frac{n_k}{2}; \frac{n_{k+1}}{2}; \phi)$.

If $\mathbf{X} = (\mathbf{X}_1', \cdots, \mathbf{X}_{k+1}')' \sim LEC_{n\times p}(0, \Sigma, \phi)$ with $\Sigma > 0$, then we say that $(\mathbf{W}_{(1)}, \cdots, \mathbf{W}_{(m)}) \sim MG_{p,k+1}(\Sigma; \frac{n_1}{2}, \cdots, \frac{n_k}{2}; \frac{n_{k+1}}{2}; \phi)$, where $\mathbf{W}_{(i)} = \mathbf{X}_i'\mathbf{X}_i$, $i = 1, \ldots, k$, $\mathbf{X}_i$, $i = 1, \ldots, k+1$, is a $n_i \times p$ matrix, $p \leq n_i < n$, $i = 1, \ldots, k$, $n_{k+1} \geq 1$, and $\sum_{i=1}^{k+1} n_i = n$.

(iii) Majorization ($\mathbf{x} \prec \mathbf{y}$).

For $\mathbf{x}, \mathbf{y} \in \Re^n$, $\mathbf{x} \prec \mathbf{y}$ ($\mathbf{y}$ majorizes $\mathbf{x}$), if

$$\begin{cases} \sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}, & i = 1, \ldots, n-1; \\ \sum_{i=1}^n x_{[i]} = \sum_{i=1}^n y_{[i]}, \end{cases} \tag{1.9}$$

where $x_{[i]}$, $i = 1, \ldots, n$, is the arrangement of x_i, $i = 1, \ldots, n$, such that $x_{[1]} \geq \cdots \geq x_{[n]}$.

In this paper invariant tests, sampling from the family of multivariate elliptically contoured distributions, are discussed. Chapter 2 develops some general results in finding the distributions and power functions of various test statistics. Chapter 3 discusses invariant

tests in multivariate regression models and their properties including admissibility. Chapter 4 studies an invariant test for equality between the mean vector and a specified vector, shows that the generalized T^2-test is the locally most powerful invariant test under some regularity conditions and is also the only admissible invariant test, and discusses the monotonicity of the power function of the generalized T^2-test. The invariant tests for testing various hypotheses and their properties are studied in Chapter 5. The following hypothesis tests are considered : lack of correlations among different sets, equality of covariance matrices and mean vectors, known correlation coefficient, known partial correlation coefficient, and zero multiple correlation coefficient.

Anderson and Fang (1982c) considered maximum likelihood estimation and the likelihood ratio test. In Chapter 6 we assume that $\mathbf{X}$ has a continuous probability density function, each element of its covariance matrix is finite, and the null hypothesis is scalar-invariant. It is then shown that the usual normal-theory likelihood ratio tests are exactly robust for the null case under this wider class (i.e. the likelihood ratio tests, sampling from this general class, are the same as the usual normal-theory likelihood ratio tests; and their null distributions are the same.)

Anderson and Fang (1982a) derive the distribution of a quadratic form for the central case for the family of multivariate elliptically contoured distributions. Chapter 7 evaluates its non-central distribution.

Chapter 2

GENERAL RESULTS FOR MULTIVARIATE ELLIPTICALLY CONTOURED DISTRIBUTIONS

In this chapter two general results for multivariate elliptically contoured distributions, used in later chapters, are derived. The multivariate normal distribution is a special case of a multivariate elliptically contoured distribution, but the distribution of a certain class of specified functions of the multivariate normal distribution is the same as that of the multivariate elliptically contoured distribution. Based on this fact, if the test statistic is one of the above specified functions, the power of this test remains the same as in the multivariate normal case.

2.1. The distributions of certain specified statistics.

Anderson and Fang (1982b, 1982c) proved that the null distributions of the likelihood ratio statistics for testing lack of correlation between sets of variates, equality between the mean vector and a specified vector, equality of covariance matrices, equality of several means, and MANOVA in the family of multivariate elliptically contoured distributions have the same distributions as in the multivariate normal case. Chmielewski (1980) also showed that the null and non-null distributions of all invariant statistics for testing the sphericity hypothesis are the same as in the multivariate normal case. In this section a general result is given.

Theorem 2.1.

Let $\mathbf{X}_{n\times p} \sim MEC_{n\times p}(\mathbf{M}; \Sigma_1, \ldots, \Sigma_1, \Sigma_2, \ldots, \Sigma_2, \ldots, \Sigma_q, \ldots, \Sigma_q; \phi)$ *with* n_i Σ_i*'s,* $\Sigma_i > 0$*, the rows of* $\mathbf{M}$ *are* n_i μ_i'*'s,* $i = 1, \ldots, q$*,* $\sum_{i=1}^{q} n_i = n$*, and the density of* $\mathbf{X}$ *exists.*

(i) If $\mathbf{M} = \mathbf{0}$ *and if the vector* $\mathbf{f}(\mathbf{X})$ *of statistics satisfies the condition that* $\mathbf{f}(\mathbf{X}) = \mathbf{f}(c\mathbf{X})$*, for every* $c > 0$*, then the distribution of* $\mathbf{f}(\mathbf{X})$ *does not depend on* ϕ*.*

(ii) If the vector $\mathbf{f}(\mathbf{X})$ *of statistics satifies the condition that* $\mathbf{f}(\mathbf{X}) = \mathbf{f}(c(\mathbf{X}+\mathbf{B}))$*, where the rows of* $\mathbf{B}$ *are* n_i $\mathbf{b}_i'$*'s, for every* $c > 0$ *and* $\mathbf{b}_i \in \Re^p$*,* $i = 1, \ldots, q$*, then the distribution of* $\mathbf{f}(\mathbf{X})$ *does not depend on* ϕ*.*

Proof.

(i) We can write

$$\mathbf{x} = \operatorname{vec}\mathbf{X}' \stackrel{d}{=} R\mathbf{D}'\mathbf{u}^{(np)}, \tag{2.1}$$

where $\mathbf{u}^{(np)}$ and R are independent, $\mathbf{D}_i'\mathbf{D}_i = \Sigma_i > 0$, and $\mathbf{D}$ is a matrix with n_1 $\mathbf{D}_1$'s, $\ldots, n_q$ $\mathbf{D}_q$'s as diagonal blocks and off-diagonal blocks of $\mathbf{0}$'s.

Define $\mathbf{f}^*(\operatorname{vec}\mathbf{X}') = \mathbf{f}(\mathbf{X})$, then $\mathbf{f}^*(\mathbf{x}) = \mathbf{f}^*(c\mathbf{x})$ and $\mathbf{f}^*(R\mathbf{D}'\mathbf{u}^{(np)}) \stackrel{d}{=} \mathbf{f}^*(cR\mathbf{D}'\,\mathbf{u}^{(np)})$ for every $c > 0$. Since the density of $\mathbf{X}$ exists, $\Pr(R = 0) = 0$. So $\mathbf{f}^*(r\mathbf{D}'\mathbf{u}^{(np)}) \stackrel{d}{=} \mathbf{f}^*(cr\mathbf{D}'\mathbf{u}^{(np)})$, where $\mathbf{f}^*(r\mathbf{D}'\mathbf{u}^{(np)})$ is the conditional random vector at $R = r > 0$. Then by letting $c = r^{-1}$, $\mathbf{f}^*(r\mathbf{D}'\mathbf{u}^{(np)}) \stackrel{d}{=} \mathbf{f}^*(\mathbf{D}'\mathbf{u}^{(np)})$. Since $\mathbf{u}^{(np)}$ and R are independent, $\mathbf{f}^*(\mathbf{x})$ is independent of R and $\mathbf{f}(\mathbf{X})$ is independent of R. So the distribution of $\mathbf{f}(\mathbf{X})$ does not depend on ϕ.

(ii) Let $\mathbf{B} = -\mathbf{M}$ and $\mathbf{Y} = \mathbf{X} - \mathbf{M}$, then

$$\begin{aligned} \mathbf{f}(\mathbf{X}) &= \mathbf{f}(c(\mathbf{X} - \mathbf{M})) \\ &= \mathbf{f}(c\mathbf{Y}). \end{aligned} \tag{2.2}$$

When $c = 1$, $\mathbf{f}(\mathbf{X}) = \mathbf{f}(\mathbf{Y})$. So it is true that $\mathbf{f}(\mathbf{Y}) = \mathbf{f}(c\mathbf{Y})$, for any $c > 0$. Then from (i), the distribution of $\mathbf{f}(\mathbf{Y})$ does not depend on ϕ. So the distribution of $\mathbf{f}(\mathbf{X})$ does not depend on ϕ. ∎

Corollary 2.1.

Let $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$ *with* $\Sigma > 0$*, and the density of* $\mathbf{X}$ *exists.*

(i) If $\mu = \mathbf{0}$ *and if the vector* $\mathbf{f}(\mathbf{X})$ *of statistics satisfies the condition that* $\mathbf{f}(\mathbf{X}) = \mathbf{f}(c\mathbf{X})$*, for every* $c > 0$*, then the distribution of* $\mathbf{f}(\mathbf{X})$ *does not depend on* ϕ*.*

(ii) If the vector $\mathbf{f}(\mathbf{X})$ *of statistics satisfies the condition that* $\mathbf{f}(\mathbf{X}) = \mathbf{f}(c(\mathbf{X} + \mathbf{1}_n\mathbf{d}'))$*, for every* $c > 0$*, and* $\mathbf{d} \in \Re^p$*, then the distribution of* $\mathbf{f}(\mathbf{X})$ *does not depend on* ϕ*.*

Proof.

It is clear by letting $q = 1$, $\Sigma_1 = \Sigma$, $\mu_1 = \mu$ in Theorem 2.1. ∎

The results of Anderson and Fang (1982b, 1982c) and Chmielewski (1980) are easily obtained from Corollary 2.1. There are a number of other examples in Chapters 3, 4, 5, and 6.

2.2. Power functions of specified tests.

The previous section consists of general results in distribution theory. In this section, we will attack the problem of power functions by using the above result.

Theorem 2.2.

(i) Assume $\mathbf{X}_{n\times p} \sim MEC_{n\times p}(\mathbf{M}; \Sigma_1, \ldots, \Sigma_1, \Sigma_2, \ldots, \Sigma_2, \ldots, \Sigma_q, \ldots, \Sigma_q; \phi)$ *with* n_i Σ_i*'s,* $\Sigma_i > 0$*, the rows of* $\mathbf{M}$ *are* n_i μ_i'*'s,* $i = 1, \ldots, q$*, and the density of* $\mathbf{X}$ *exists. For testing* $H_0 : (\mathbf{M}, \Sigma_1, \ldots, \Sigma_q) \in \Omega_0$ *vs* $H_1 : (\mathbf{M}, \Sigma_1, \ldots, \Sigma_q) \in \Omega \setminus \Omega_0$*, if the hypotheses remain invariant under the group* $G = \{g \mid g : \mathbf{X} \mapsto c(\mathbf{X} + \mathbf{B})$*, where the rows of* $\mathbf{B}$ *are* n_i $\mathbf{b}_i'$*'s,* $c > 0$*,* $\mathbf{b}_i \in \Re^p$*,* $i = 1, \ldots, q\}$*, and* $f(\mathbf{X})$ *is an invariant test statistic, then the power function of this test with rejection region* $f(\mathbf{X}) \in S_1$ *does not depend on* ϕ*.*

(ii) $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$*, with* $\Sigma > 0$*, and the density of* $\mathbf{X}$ *exists. For testing* $H_0 : (\mu, \Sigma) \in \Omega_0$ *vs* $H_1 : (\mu, \Sigma) \in \Omega \setminus \Omega_0$*, if the hypotheses remain invariant under the group*

$\{g \mid g : \mathbf{X} \mapsto c(\mathbf{X} + \mathbf{1}_n\mathbf{d}'),\ c > 0,\ \mathbf{d} \in \Re^p\}$, *and* $f(\mathbf{X})$ *is an invariant test statistic, then the power function of this test with rejection region* $f(\mathbf{X}) \in S_1$ *does not depend on* ϕ.

Proof.

(i) Since $f(\mathbf{X}) = f(c(\mathbf{X} + \mathbf{B}))$, by Theorem 2.1, the distribution of $f(\mathbf{X})$ does not depend on ϕ under both the null and the non-null cases. Therefore the power function of this invariant test does not depend on ϕ.

(ii) This is a special case of (i). Using this, we can complete the proof. ∎

Corollary 2.2.

(i) In Theorem 2.2 (i), if $\Omega = B \times \Omega^*$, $\Omega_0 = B \times \Omega_0^*$, $\Omega \setminus \Omega_0 = B \times (\Omega^* \setminus \Omega_0^*)$, *where* B *is the set of parameters* $\mathbf{M}$, Ω^* *is the set of parameters* $(\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_q)$, *and* $\Omega_0^* \subset \Omega^*$, $f(\mathbf{X})$ *is an invariant test statistic of* $H_0 : (\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_q) \in \Omega_0^*$ *vs* $H_1 : (\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_q) \in \Omega^* \setminus \Omega_0^*$, *then the power function of this test with same rejection region* S_1 *does not depend on* ϕ.

(ii) In Theorem 2.2 (ii), if $\Omega = B \times \Omega^*$, $\Omega_0 = B \times \Omega_0^*$, $\Omega \setminus \Omega_0 = B \times (\Omega^* \setminus \Omega_0^*)$, *where* B *is the set of parameters* μ , Ω^* *is the set of parameters* $\mathbf{\Sigma}$, *and* $\Omega_0^* \subset \Omega^*$, $f(\mathbf{X})$ *is an invariant test statistic of* $H_0 : \mathbf{\Sigma} \in \Omega_0^*$ *vs* $H_1 : \mathbf{\Sigma} \in \Omega^* \setminus \Omega_0^*$, *then the power function of this test with same rejection region* S_1 *does not depend on* ϕ.

Proof.

(i) It is clear that the hypotheses remain invariant under the group $\{g \mid g : \mathbf{X} \mapsto \mathbf{X} + \mathbf{B}$ where the rows of $\mathbf{B}$ are n_i $\mathbf{b}_i'$'s, $\mathbf{b}_i \in \Re^p$, $i = 1, \ldots, q\}$. Since under the transformation $\mathbf{X} \mapsto c\mathbf{X}$, $c > 0$, $(\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_q)$ does not change (only ϕ changes) (Since $\mathbf{\Sigma}_i$, $i = 1, \ldots, q$, are scale matrix. i.e., for any $a > 0$, $\mathbf{X} \sim MEC_{n \times p}(0; \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \ldots, \mathbf{\Sigma}_2, \ldots, \mathbf{\Sigma}_q, \ldots, \mathbf{\Sigma}_q; \phi)$,then $a\mathbf{X} \sim MEC_{n \times p}(0; \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \ldots, \mathbf{\Sigma}_2, \ldots, \mathbf{\Sigma}_q, \ldots, \mathbf{\Sigma}_q; \phi^*)$, see Cambanis, Huang and Simons (1981).) So by Theorem 2.2, the power function of this test with same rejection region S_1 does not depend on ϕ.

(ii) Similarly, as in (i), we can easily complete the proof. ∎

We will apply this result frequently in the later chapters. In particular, under the same assumptions as Corollary 2.2, if an invariant test for the multivariate elliptically contoured distribution is UMPI (uniformly most powerful invariant test) in the multivariate normal case, then we can claim that this test is also UMPI in the multivariate elliptically contoured distribution.

Chapter 3

MULTIVARIATE REGRESSION ANALYSIS

In this chapter, some properties in a multivariate regression model are studied. The sampling is from the family of multivariate elliptically contoured distributions. Invariant tests are derived and turn out to be the same as those dealt with in the multivariate normal case. Admissibility is a desirable property. Admissibility and null distributions are studied.

3.1. Invariant tests and null distributions.

The usual assumption in the multivariate regression model is that the sample consists of independent and identically distributed normal vectors. Now we consider, the more general case, the following multivariate regression model sampling from the family of multivariate elliptically contoured distributions :

$$\begin{cases} \mathbf{Y}_{n\times p} = \mathbf{X}_{n\times q}\mathbf{B}_{q\times p} + \mathbf{E}_{n\times p}, \\ \\ \mathbf{E} \sim LEC_{n\times p}(0, \Sigma, \phi), \end{cases} \tag{3.1}$$

where $\mathbf{X}$ is known, $q \le p \le n$, $n - q \ge p$, $\operatorname{rank}(\mathbf{X}) = q$, and $\Sigma > 0$.

If the density of $\mathbf{Y}$ exists, then the density of $\mathbf{E}$ exists with the form of

$$|\Sigma|^{-\frac{1}{2}n}\, g(\operatorname{tr}\Sigma^{-1}(\mathbf{E}'\mathbf{E})). \tag{3.2}$$

Hence the density of $\mathbf{Y}$ has the form of

$$|\Sigma|^{-\frac{1}{2}n}\, g(\operatorname{tr}\Sigma^{-1}(\mathbf{Y}-\mathbf{XB})'(\mathbf{Y}-\mathbf{XB})). \tag{3.3}$$

Suppose we partition

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix}, \qquad \mathbf{X} = (\mathbf{X}_1 \quad \mathbf{X}_2), \tag{3.4}$$

such that $\mathbf{B}_1$ has q_1 rows, $\mathbf{B}_2$ has q_2 rows, $\mathbf{X}_1$ has q_1 columns, and $\mathbf{X}_2$ has q_2 columns. We wish to test $H : \mathbf{B}_1 = \mathbf{B}_1^*$ against the general alternative $K : \mathbf{B}_1 \neq \mathbf{B}_1^*$, where $\mathbf{B}_1^*$ is a given matrix.

To find the maximal invariants of the sufficient statistics, we need to find the sufficient statistics. Since

$$\mathbf{Y} - \mathbf{X}\mathbf{B} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}_\Omega) + \mathbf{X}(\hat{\mathbf{B}}_\Omega - \mathbf{B}), \tag{3.5}$$

we have

$$\begin{aligned}(\mathbf{Y} - \mathbf{X}\mathbf{B})'(\mathbf{Y} - \mathbf{X}\mathbf{B}) \\ &= (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}_\Omega)'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}_\Omega) + (\hat{\mathbf{B}}_\Omega - \mathbf{B})'\mathbf{X}'\mathbf{X}(\hat{\mathbf{B}}_\Omega - \mathbf{B}) \\ &= N\hat{\boldsymbol{\Sigma}}_\Omega + (\hat{\mathbf{B}}_\Omega - \mathbf{B})'\mathbf{A}(\hat{\mathbf{B}}_\Omega - \mathbf{B}),\end{aligned} \tag{3.6}$$

where $N\hat{\boldsymbol{\Sigma}}_\Omega = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}_\Omega)'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}_\Omega)$, $\quad \hat{\mathbf{B}}_\Omega = \mathbf{A}^{-1}\mathbf{C}$, $\quad \mathbf{C} = \mathbf{X}'\mathbf{Y}$, $\quad$ and

$$\mathbf{A} = \mathbf{X}'\mathbf{X} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}.$$

From (3.3), the density of $\mathbf{Y}$ has the form of

$$|\boldsymbol{\Sigma}|^{-\frac{1}{2}n}\, g(\operatorname{tr} \boldsymbol{\Sigma}^{-1}[N\hat{\boldsymbol{\Sigma}}_\Omega + (\hat{\mathbf{B}}_\Omega - \mathbf{B})'\mathbf{A}(\hat{\mathbf{B}}_\Omega - \mathbf{B})]). \tag{3.7}$$

By the Factorization theorem, $\hat{\boldsymbol{\Sigma}}_\Omega$, $\hat{\mathbf{B}}_\Omega$ form a sufficient set of statistics for $\boldsymbol{\Sigma}$, $\mathbf{B}$. Thus, there is an one-to-one correspondence between $(\hat{\mathbf{B}}_{1\Omega}, \hat{\mathbf{B}}_{2\omega})$ and $(\hat{\mathbf{B}}_{1\Omega}, \hat{\mathbf{B}}_{2\Omega})$ and $\hat{\boldsymbol{\Sigma}}_\Omega$, $\hat{\mathbf{B}}_{1\Omega}$, and $\hat{\mathbf{B}}_{2\omega}$ form a sufficient set of statistics for $\boldsymbol{\Sigma}$, $\mathbf{B}$. (Note that $\hat{\mathbf{B}}_{2\omega} = \hat{\mathbf{B}}_{2\Omega} + (\hat{\mathbf{B}}_{1\Omega} - \mathbf{B}_1^*)\mathbf{A}_{12}\mathbf{A}_{22}^{-1}$, where $\hat{\mathbf{B}}_{2\omega} = \mathbf{A}_{22}^{-1}[\mathbf{X}_2'(\mathbf{Y} - \mathbf{X}_1\mathbf{B}_1^*)]$.)

We can reformulate the hypothesis as $\mathbf{B}_1 = \mathbf{0}$ (by replacing $\mathbf{Y}$ by $\mathbf{Y} - \mathbf{X}_1\mathbf{B}_1^*$) and the

problem remains invariant under

$$(1)\mathbf{Y} \mapsto \mathbf{Y} + \mathbf{X}_2\mathbf{\Gamma},$$

$$(2)\mathbf{X}_1 \mapsto \mathbf{X}_1\mathbf{C}^*, \qquad \mathbf{B}_1 \mapsto \mathbf{C}^{*-1}\mathbf{B}_1, \tag{3.8}$$

where $\mathbf{C}^*$ is nonsingular,

$$(3)\mathbf{Y} \mapsto \mathbf{YV},$$

where $\mathbf{V}$ is nonsingular.

Under group(1),

$$\begin{aligned} &\hat{\mathbf{\Sigma}}_\Omega \mapsto \hat{\mathbf{\Sigma}}_\Omega, \qquad \hat{\mathbf{B}}_{1\Omega} \mapsto \hat{\mathbf{B}}_{1\Omega}, \\ &\hat{\mathbf{B}}_{2\omega} \mapsto \hat{\mathbf{B}}_{2\omega} + \mathbf{\Gamma}. \end{aligned} \tag{3.9}$$

So the only invariants of the sufficient statistics are $\hat{\mathbf{\Sigma}}_\Omega$ and $\hat{\mathbf{B}}_{1\Omega}$ (by letting $\mathbf{\Gamma} = -\hat{\mathbf{B}}_{2\omega}$). Under group (2),

$$\begin{aligned} &\hat{\mathbf{\Sigma}}_\Omega \mapsto \hat{\mathbf{\Sigma}}_\Omega, \qquad \hat{\mathbf{B}}_{1\Omega} \mapsto \mathbf{C}^{*-1}\hat{\mathbf{B}}_{1\Omega}, \\ &\mathbf{X}_1^* \mapsto \mathbf{X}_1^*\mathbf{C}^*, \end{aligned} \tag{3.10}$$

where $\mathbf{X}_1^* = \mathbf{X}_1 - \mathbf{X}_2\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$. Let $\mathbf{C}_1^*\mathbf{A}_{11\cdot 2}\mathbf{C}_1^* = \mathbf{I}$ and $\mathbf{C}_2^*$ be an orthogonal transformation such that

$$\mathbf{C}_2^{*-1}\mathbf{C}_1^{*-1}\hat{\mathbf{B}}_{1\Omega} = \mathbf{T}, \tag{3.11}$$

where $t_{iv} = 0$, $i > v$, $t_{ii} \geq 0$, and $\mathbf{A}_{11\cdot 2} = \mathbf{X}_1^{*\prime}\mathbf{X}_1^* = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}{}^{-1}\mathbf{A}_{21}$. Since $\mathbf{T}$ is a function of $\mathbf{T}'\mathbf{T} = \hat{\mathbf{B}}_{1\Omega}'\mathbf{A}_{11\cdot 2}\hat{\mathbf{B}}_{1\Omega}$, the only invariants of the sufficient statistics are $N\hat{\mathbf{\Sigma}}_\Omega$ (denote as $\mathbf{G}$) and $\hat{\mathbf{B}}_{1\Omega}'\mathbf{A}_{11\cdot 2}\hat{\mathbf{B}}_{1\Omega}$ (denote as $\mathbf{H}$). Under group (3),

$$\mathbf{G} \mapsto \mathbf{V}'\mathbf{GV}, \qquad \mathbf{H} \mapsto \mathbf{V}'\mathbf{HV}, \tag{3.12}$$

where $\mathbf{V}$ is nonsingular. Hence the only invariants (i.e. maximal invariants) of the sufficient statistics are the roots of

$$|\mathbf{H} - \theta\mathbf{G}| = 0. \tag{3.13}$$

We have the following theorem.

Theorem 3.1.

The respective maximal invariants of the sufficient statistics for the sample space are the roots of (3.13).

Since the maximal invariants of the sufficient statistics are same as the ones in multivariate normal case, the invariant tests are same as the ones in multivariate normal case. Let $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_p$ be the roots of (3.13). Some invariant test criteria are listed as follows:

(1) Lawley-Hotelling trace:

$$\operatorname{tr}\mathbf{H}\mathbf{G}^{-1} = \sum_{i=1}^{p} \theta_i. \tag{3.14}$$

(2) Bartlett-Nanda-Pillai trace:

$$\operatorname{tr}\mathbf{H}(\mathbf{H}+\mathbf{G})^{-1} = \sum_{i=1}^{p} \frac{\theta_i}{1+\theta_i}. \tag{3.15}$$

(3) Roy's maximum root:

$$\theta_1. \tag{3.16}$$

(4) Wilks' likelihood ratio criterion:

$$U = \frac{|\mathbf{G}|}{|\mathbf{G}+\mathbf{H}|} = \prod_{i=1}^{p}(1+\theta_i)^{-1}. \tag{3.17}$$

Under the null hypothesis, by Theorem 2.1, the joint distribution of maximal invariants of the sufficient statistics does not depend on ϕ. So the null distribution of an invariant test is the same for all ϕ's (i.e. it is same as the one in multivariate normal case.)

Anderson and Fang (1982b) derived that $N\hat{\Sigma}_\Omega$ is distributed as $MG_{p,2}(\Sigma; \frac{n-q}{2}; \frac{q}{2}; \phi)$ (see page 4.) Next we want to find the null distribution of $\mathbf{H}$. Since

$$\hat{\mathbf{B}}_{1\Omega} = \mathbf{A}^{(1)}\mathbf{X}'\mathbf{E}, \tag{3.18}$$

and

$$\mathbf{E} \stackrel{d}{=} R\mathbf{U}_{n\times p}\mathbf{D}_{p\times p}, \tag{3.19}$$

$$\begin{aligned} \mathbf{H} &= \hat{\mathbf{B}}_{1\Omega}\mathbf{A}_{11\cdot 2}\hat{\mathbf{B}}_{1\Omega} \\ &= \mathbf{E}'\mathbf{X}\mathbf{A}^{(1)'}\mathbf{A}_{11.2}\mathbf{A}^{(1)}\mathbf{X}'\mathbf{E} \\ &= \mathbf{E}'\mathbf{F}\mathbf{E}, \end{aligned} \tag{3.20}$$

where $\mathbf{F} = \mathbf{X}\mathbf{A}^{(1)'}\mathbf{A}_{11\cdot 2}\mathbf{A}^{(1)}\mathbf{X}'$, $\quad \mathbf{D}'\mathbf{D} = \Sigma > 0$, $\quad$ and

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \end{pmatrix}.$$

If $R^2 \sim \chi^2_{np}$ (i.e., $\mathbf{E}$ has a multivariate normal distribution), then $\mathbf{H} = \mathbf{E}'\mathbf{F}\mathbf{E} \sim W_p(\Sigma, q_1)$. From Cochran's Theorem for multivariate normal distributions, we have $\mathbf{F}^2 = \mathbf{F}$ and $\mathrm{rank}(\mathbf{F}) = q_1$. Then by Cochran's Theorem for multivariate elliptically contoured distributions (Anderson and Fang (1982b)), $\mathbf{H} \sim MG_{p,2}(\Sigma; \frac{q_1}{2}; \frac{n-q_1}{2}; \phi)$.

3.2. Admissibility of invariant tests.

In multivariate normality, one way to approach admissibility is to apply either Stein's Theorem or Schwartz's Theorem (Section 8.10 of Anderson (1984)). In this section, we discuss the problem of admissibility and make an extension of Schwartz's Theorem to multivariate elliptically contoured distributions.

Theorem 3.2.

Assume that $(\mathcal{Y}, \mathrm{m}, \Omega^*, P^*)$ is a family of distributions and $(\mathcal{Y}, \mathrm{m}, \Omega, P)$ is a subfamily of $(\mathcal{Y}, \mathrm{m}, \Omega^*, P^*)$ such that for any bounded function f, $E_\omega f(Y) \equiv 0$ $\forall \omega \in \Omega$ implies $f(Y) \equiv 0$ a.e. m. Suppose Ω_0^* and Ω_0 are the nonempty proper subsets of Ω^* and Ω, respectively,

such that $\Omega_0 \subset \Omega_0^*$ and $\Omega \setminus \Omega_0 \subset \Omega^* \setminus \Omega_0^*$. If the test with acceptance region B is admissible for testing the hypothesis that $\omega \in \Omega_0$ against the alternative $\omega \in \Omega \setminus \Omega_0$, then the test with acceptance region B is admissible for testing the hypothesis that $\omega \in \Omega_0^*$ against the alternative $\omega \in \Omega^* \setminus \Omega_0^*$.

Proof.

The critical function of the test with acceptance region B is $\phi_B(y) = 0$, $y \in B$, and $\phi_B(y) = 1$, $y \notin B$. Suppose $\phi(y)$ is the critical function of a better test, that is

$$\begin{aligned} \int \phi(y)\, dp_\omega^*(y) &\leq \int \phi_B(y)\, dp_\omega^*(y), \qquad \omega \in \Omega_0^*, \\ \int \phi(y)\, dp_\omega^*(y) &\geq \int \phi_B(y)\, dp_\omega^*(y), \qquad \omega \in \Omega^* \setminus \Omega_0^*, \end{aligned} \tag{3.21}$$

with strict inequality for some ω; we shall show that this assumption leads to a contradiction.

Now we restrict the above inequalities to Ω; we have

$$\begin{aligned} \int (\phi(y) - \phi_B(y))\, dp_\omega(y) &\leq 0, \qquad \omega \in \Omega_0, \\ \int (\phi(y) - \phi_B(y))\, dp_\omega(y) &\geq 0, \qquad \omega \in \Omega \setminus \Omega_0. \end{aligned} \tag{3.22}$$

Since the test with acceptance region B is admissible for testing the hypothesis that $\omega \in \Omega_0$ against the alternative $\omega \in \Omega \setminus \Omega_0$,

$$E_\omega[\phi(Y) - \phi_B(Y)] \equiv 0 \qquad \forall \omega \in \Omega. \tag{3.23}$$

So $\phi(Y) \equiv \phi_B(Y)$ a.e. m. This leads to a contradiction. ∎

Now we go back to the problem of the multivariate regression model. Since $p \geq q_1$, there are $p - q_1$ roots of (3.13) identically 0. It seems reasonable that if a set of roots $\theta = (\theta_1, \ldots, \theta_{q_1})'$ leads to acceptance, then a set of roots $\theta^* = (\theta_1^*, \ldots, \theta_{q_1}^*)'$ such that $\theta_i^* \leq \theta_i$, $i = 1, \ldots, q_1$, should also lead to acceptance. Such an acceptance set is called monotone. A

point $\theta = (\theta_1, \ldots, \theta_{q_1})'$ is in the extended region A^* if the point $\theta_p = (\theta_{i(1)}, \ldots, \theta_{i(q_1)})'$, where $i(1), \ldots, i(q_1)$ is a permutation of $1, \ldots, q_1$ such that $\theta_{i(1)} \geq \cdots \geq \theta_{i(q_1)}$, is in A. The following theorem is a generalization of Schwartz's Theorem.

Theorem 3.3.

Let $(\mathcal{Y}, \mathrm{m}, \Omega^*, \mathcal{P}^*)$ be the family of multivariate elliptically contoured distributions and m be the Lebesgue measure. If the acceptance region A in the set $1 \geq \theta_1 \geq \cdots \geq \theta_{q_1} \geq 0$ is montone and if the extended region A^* is closed and convex, then A is the acceptance region of an admissible test.

This theorem is proved using the following corollary from Theorem 3.2, and Theorem 3.4.1 of Ferguson (1967).

Corollary 3.1.

Let $(\mathcal{Y}, \mathrm{m}, \Omega^*, \mathcal{P}^*)$ be the family of multivariate elliptically contoured distributions and m be the Lebesgue measure. If the acceptance region A in the set $1 \geq \theta_1 \geq \cdots \geq \theta_{q_1} \geq 0$ is montone and if the extended region A^* is closed and convex, then A is the acceptance region of an admissible test among the class of tests based on the sufficient statistics $(\hat{\Sigma}_\Omega, \hat{\mathbf{B}}_\Omega)$.

Proof.

Let $(\mathcal{Y}, \mathrm{m}, \Omega, \mathcal{P})$ be the family of multivariate normal distributions. It is clear that $(\hat{\Sigma}_\Omega, \hat{\mathbf{B}}_\Omega)$ is the complete sufficient statistics. Then this corollary follows from Theorem 3.2. ∎

Next let us state Theorem 3.4.1 of Ferguson (1967).

Theorem 3.4.

Consider the game $(\Theta, \mathcal{A}, L)$ where the statistician observes a random vector $\mathbf{X}$ whose distribution depends on θ. If T is a sufficient statistic for θ, then the set $\mathcal{D}_0$ of decision

rules in D, which are based on T, forms an essentially complete class in the game (Θ, D, $\hat{R}$).

Note :

Θ is the parameter space; A is the space of actions; L is a loss function; D is the space of all behavioral decision rules; and $\hat{R}$ is the risk function (see Ferguson (1967).)

Proof of Theorem *3.3.*

For every fixed density g (see (3.2)), from Corollary 3.1, A is the acceptance region of an admissible test among the class of tests based on the sufficient statistic $(\hat{\mathbf{\Sigma}}_\Omega, \hat{\mathbf{B}}_\Omega)$. But $(\hat{\mathbf{\Sigma}}_\Omega, \hat{\mathbf{B}}_\Omega)$ is the sufficient statistic for $(\mathbf{\Sigma}, \mathbf{B})$. Then, from Theorem 3.4, A is the acceptance region of an admissible test for this fixed g. So A is the acceptance region of an admissible test. ∎

From Theorem 3.3, we conclude that Wilks' likelihood ratio test, the Lawley-Hotelling trace test, the Bartlett-Nanda-Pillai trace test, and Roy's maximum root test are admissible.

Chapter 4

TESTING HYPOTHESIS FOR MEAN VECTOR

In this chapter an invariant test for equality between the mean vector and a specified vector is derived; other optimum properties derived in Chapter 3 are studied. In the multivariate normal case, the generalized T^2-test is the uniformly most powerful invariant test (UMPI), but in the multivariate elliptically contoured case, a weaker property called locally most powerful invariant test (LMPI) is derived in section 4.2.

4.1. Invariant test, null distribution, and admissibility.

Suppose $\mathbf{X}_{n \times p} \sim LEC_{n \times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$, unknown, $n \geq p$, and the density of $\mathbf{X}$ exists. We wish to test $H : \mu = \mu_0$ against the general alternative $K : \mu \neq \mu_0$, where μ_0 is a known vector. We can reformulate the hypothesis as $\mu = \mathbf{0}$ (by replacing $\mathbf{X}$ by $\mathbf{X} - \mathbf{1}_n \mu_0'$) and the problem remains invariant under the following group :

$$\mathbf{X} \mapsto \mathbf{XC}, \tag{4.1}$$

where $\mathbf{C}$ is $p \times p$ nonsingular.

This problem is a special case of multivariate regression model, so we have the following corollaries.

Corollary 4.1.

The respective maximal invariant of the sufficient statistics for the sample space is $T^2 = n(\bar{\mathbf{x}} - \mu_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \mu_0)$ *where* $\bar{\mathbf{x}} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n$ *and* $\mathbf{S} = \frac{1}{n-1}(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')$.

Corollary 4.1 implies that the T^2-test whose acceptance region is $T^2 \leq \lambda_1$, a constant, is an invariant test.

Corollary 4.2.

For the null case, the distribution of T^2 *does not depend on the distribution of the radius* R. *That is, the null distribution is unique for any* ϕ.

Corollary 4.3.

T^2*-test is admissible.*

4.2. Locally most powerful invariant test.

In this section, we want to derive a nice result that the T^2-test is a locally most powerful invariant test. First we need to find the non-null distribution of T^2.

If $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$, $n \geq p > 1$, and the density of $\mathbf{X}$ exists, then

$$\mathbf{X} \stackrel{d}{=} \mathbf{1}_n\mu' + R\mathbf{U}_{n\times p}\mathbf{D}_{p\times p}, \tag{4.2}$$

and

$$\begin{aligned} \bar{\mathbf{x}} &= \frac{1}{n}\mathbf{X}'\mathbf{1}_n \\ &\stackrel{d}{=} \mu + \frac{1}{n}R\mathbf{D}'\mathbf{U}'\mathbf{1}_n, \end{aligned} \tag{4.3}$$

where $\mathbf{D}'\mathbf{D} = \Sigma$, $R \sim F$ and independent of $\mathbf{U}$. So

$$\frac{T^2}{n-1} = \frac{n}{n-1}(\bar{\mathbf{x}}-\mu_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}}-\mu_0)$$

$$\stackrel{d}{=} n[(\mu-\mu_0)+\frac{1}{n}R\mathbf{D}'\mathbf{U}'\mathbf{1}_n]'[R^2\mathbf{D}'\mathbf{U}'(\mathbf{I}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\mathbf{U}\mathbf{D}]^{-1}$$

$$\times\,[(\mu-\mu_0)+\frac{1}{n}R\mathbf{D}'\mathbf{U}'\mathbf{1}_n]$$

$$= [\frac{n^{\frac{1}{2}}}{R}(\mathbf{D}^{-1})'(\mu-\mu_0)+\frac{1}{n^{\frac{1}{2}}}\mathbf{U}'\mathbf{1}_n]'[\mathbf{U}'(\mathbf{I}-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\mathbf{U}]^{-1}$$

$$\times\,[\frac{n^{\frac{1}{2}}}{R}(\mathbf{D}^{-1})'(\mu-\mu_0)+\frac{1}{n^{\frac{1}{2}}}\mathbf{U}'\mathbf{1}_n]. \tag{4.4}$$

The last equality of (4.4) is true because of $\Pr\{R=0\}=0$.

Let

$$\mathbf{Y} = \frac{n^{\frac{1}{2}}}{R}(\mathbf{D}^{-1})'(\mu-\mu_0)+\frac{1}{n^{\frac{1}{2}}}\mathbf{U}'\mathbf{1}_n, \tag{4.5}$$

and $\mathbf{Q}$ be a $p\times p$ orthogonal matrix with the first row $\frac{1}{(\mathbf{Y}'\mathbf{Y})^{\frac{1}{2}}}\mathbf{Y}'$. Define

$$\mathbf{V} = \mathbf{QY}$$

$$= \begin{pmatrix} (\mathbf{Y}'\mathbf{Y})^{\frac{1}{2}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{4.6}$$

and

$$\begin{aligned} \mathbf{B} &= \mathbf{Q}[\mathbf{U}'(\mathbf{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\mathbf{U}]\mathbf{Q}' \\ &= \begin{pmatrix} b_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}, \end{aligned} \tag{4.7}$$

then

$$\begin{aligned} \frac{T^2}{n-1} &\stackrel{d}{=} \mathbf{V}'\mathbf{B}^{-1}\mathbf{V} \\ &= \mathbf{Y}'\mathbf{Y}b^{11} \\ &= \frac{\mathbf{Y}'\mathbf{Y}}{b_{11\cdot 2}}, \end{aligned} \tag{4.8}$$

where b^{11} is the (1,1)th element of $\mathbf{B}^{-1}$ and $b_{11\cdot 2} = b_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21}$.

Let $\mathbf{E}$ be a $n \times n$ orthogonal matrix with nth row $\frac{1}{n^{\frac{1}{2}}}\mathbf{1}_n'$ and

$$\mathbf{EU} = \begin{pmatrix} \mathbf{u}_1' \\ \vdots \\ \mathbf{u}_n' \end{pmatrix} \stackrel{d}{=} \mathbf{U}, \tag{4.9}$$

then

$$\begin{aligned} \mathbf{B} &= \mathbf{QU}'\mathbf{UQ}' - \frac{1}{n}\mathbf{QU}'\mathbf{1}_n\mathbf{1}_n'\mathbf{UQ}' \\ &= \mathbf{QU}'(\mathbf{E}'\mathbf{E})\mathbf{UQ}' - \frac{1}{n}\mathbf{QU}'\mathbf{1}_n\mathbf{1}_n'\mathbf{UQ}' \\ &= \mathbf{Q}(\sum_{i=1}^{n}\mathbf{u}_i\mathbf{u}_i')\mathbf{Q}' - \mathbf{Qu}_n\mathbf{u}_n'\mathbf{Q}' \\ &= \mathbf{Q}(\sum_{i=1}^{n-1}\mathbf{u}_i\mathbf{u}_i')\mathbf{Q}', \end{aligned} \tag{4.10}$$

and

$$\mathbf{Y} = \frac{n^{\frac{1}{2}}}{R}(\mathbf{D}^{-1})'(\mu - \mu_0) + \mathbf{u}_n. \tag{4.11}$$

From lemma 2 due to Cambanis, Huang and Simons (1981),

$$\mathbf{U} \stackrel{d}{=} \begin{pmatrix} (1-R_1^2)^{\frac{1}{2}}\mathbf{U}_1 \\ R_1\mathbf{u}^{*\prime} \end{pmatrix}, \tag{4.12}$$

where R_1 (≥ 0), $\mathbf{U}_1$ ($(n-1) \times p$) and $\mathbf{u}^*$ ($p \times 1$) are independent, and $R_1^2 \sim \text{Beta}(\frac{p}{2}, \frac{(n-1)p}{2})$. Thus

$$\begin{aligned} \mathbf{B} &\stackrel{d}{=} (1-R_1^2)\mathbf{Q}\mathbf{U}_1'\mathbf{U}_1\mathbf{Q}' \\ &\equiv \mathbf{B}(R, R_1, \mathbf{U}_1, \mathbf{u}^*), \end{aligned} \tag{4.13}$$

and

$$\begin{aligned} \mathbf{Y} &\stackrel{d}{=} \frac{n^{\frac{1}{2}}}{R}(\mathbf{D}^{-1})'(\mu - \mu_0) + R_1\mathbf{u}^* \\ &\equiv \mathbf{Y}(R, R_1, \mathbf{u}^*). \end{aligned} \tag{4.14}$$

Consequently

$$\begin{aligned} \Pr\{T^2 \leq x\} &= \Pr\left\{\frac{\mathbf{Y}'(R, R_1, \mathbf{u}^*)\mathbf{Y}(R, R_1, \mathbf{u}^*)}{b_{11\cdot 2}(R, R_1, \mathbf{U}_1, \mathbf{u}^*)} \leq \frac{x}{n-1}\right\} \\ &= \int_0^\infty \int_0^1 \Pr\left\{\frac{\mathbf{Y}'(r, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(r, v^{\frac{1}{2}}, \mathbf{u}^*)}{b_{11\cdot 2}(r, v^{\frac{1}{2}}, \mathbf{U}_1, \mathbf{u}^*)} \leq \frac{x}{n-1}\right\} \\ &\quad \times \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv\, dF(r), \end{aligned} \tag{4.15}$$

where $\beta(\cdot\,; a, b)$ is the density function of Beta(a, b).

Conditioned on $\mathbf{u}^* = \mathbf{u}$ in $\mathbf{B}(r, v^{\frac{1}{2}}, \mathbf{U}_1, \mathbf{u}^*)$, $\mathbf{Q}$ is a constant orthogonal matrix. So

$$\begin{aligned} \mathbf{B}(r, v^{\frac{1}{2}}, \mathbf{U}_1, \mathbf{u}) &= (1 - v)\mathbf{Q}\mathbf{U}_1'\mathbf{U}_1\mathbf{Q}' \\ &\stackrel{d}{=} (1 - v)\mathbf{U}_1'\mathbf{U}_1, \end{aligned} \tag{4.16}$$

does not depend on $\mathbf{u}$. Hence $\mathbf{B}(r, v^{\frac{1}{2}}, \mathbf{U}_1, \mathbf{u}^*)$ and $\mathbf{Y}'(r, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(r, v^{\frac{1}{2}}, \mathbf{u}^*)$ are independent, as are $b_{11\cdot 2}(r, v^{\frac{1}{2}}, \mathbf{U}_1, \mathbf{u}^*)$ and $\mathbf{Y}'(r, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(r, v^{\frac{1}{2}}, \mathbf{u}^*)$.

Let

$$\begin{aligned} \mathbf{U}_1'\mathbf{U}_1 &= \mathbf{E} \\ &= \begin{pmatrix} e_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{pmatrix}, \end{aligned} \tag{4.17}$$

and

$$\mathbf{U}_1 = (\mathbf{u}_1^* \quad \mathbf{U}_2), \tag{4.18}$$

where $\mathbf{u}_1^*$ is $(n-1) \times 1$, $\mathbf{U}_2$ is $(n-1) \times (p-1)$, then

$$(\mathbf{u}_1^* \quad \mathbf{U}_2) \stackrel{d}{=} (R_2\mathbf{v}_1 \qquad (1 - R_2^2)^{\frac{1}{2}}\mathbf{V}_2), \tag{4.19}$$

where $R_2 (\geq 0)$, $\mathbf{v}_1$ and $\mathbf{V}_2$ are independent, and $R_2^2 \sim \text{Beta}(\frac{n-1}{2}, \frac{(n-1)(p-1)}{2})$. So

$$\begin{aligned} e_{11\cdot 2} &= e_{11} - \mathbf{E}_{12}\mathbf{E}_{22}^{-1}\mathbf{E}_{21} \\ &= \mathbf{u}_1^{*\prime}\mathbf{u}_1^* - \mathbf{u}_1^{*\prime}\mathbf{U}_2(\mathbf{U}_2'\mathbf{U}_2)^{-1}\mathbf{U}_2'\mathbf{u}_1^* \\ &\stackrel{d}{=} R_2^2\mathbf{v}_1'\mathbf{v}_1 - R_2^2\mathbf{v}_1'\mathbf{V}_2(\mathbf{V}_2'\mathbf{V}_2)^{-1}\mathbf{V}_2'\mathbf{v}_1 \\ &= R_2^2\mathbf{v}_1'(\mathbf{I} - \mathbf{V}_2(\mathbf{V}_2'\mathbf{V}_2)^{-1}\mathbf{V}_2')\mathbf{v}_1. \end{aligned} \tag{4.20}$$

Since $n \geq p > 1$, $\mathbf{V}_2'\mathbf{V}_2$ is nonsingular with probability one. $\mathbf{A} = \mathbf{I} - \mathbf{V}_2(\mathbf{V}_2'\mathbf{V}_2)^{-1}\mathbf{V}_2'$, conditioned on $\mathbf{V}_2$, is a constant symmetric matrix, $\mathbf{A} = \mathbf{A}^2$ and $\operatorname{rank}(\mathbf{A}) = (n-1)-(p-1) = n-p$. By Corollary 1 due to Anderson and Fang (1982a), $\mathbf{v}_1'\mathbf{A}\mathbf{v}_1 \sim \operatorname{Beta}(\frac{n-p}{2}, \frac{p-1}{2})$ does not depend on $\mathbf{V}_2$. So $\mathbf{v}_1'(\mathbf{I}-\mathbf{V}_2(\mathbf{V}_2'\mathbf{V}_2)^{-1}\mathbf{V}_2')\mathbf{v}_1 \sim \operatorname{Beta}(\frac{n-p}{2}, \frac{p-1}{2})$ and $e_{11\cdot 2} \sim \operatorname{Beta}(\frac{n-p}{2}, \frac{n(p-1)}{2})$. Then

$$\Pr\{T^2 \leq x\} = \int_0^\infty \int_0^1 \int_0^1 \Pr\left\{\mathbf{Y}'(r, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(r, v^{\frac{1}{2}}, \mathbf{u}^*) \leq \frac{x(1-v)w}{n-1}\right\}$$
$$\times\, \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2})\, dw\, \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv\, dF(r). \tag{4.21}$$

Let $y = xw$, then

$$\Pr\{T^2 \leq x\} = \int_0^\infty \int_0^1 \int_0^x \Pr\left\{\mathbf{Y}'(r, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(r, v^{\frac{1}{2}}, \mathbf{u}^*) \leq \frac{y(1-v)}{n-1}\right\}$$
$$\times\, \frac{\Gamma(\frac{(n-1)p}{2})}{\Gamma(\frac{n-p}{2})\Gamma(\frac{n(p-1)}{2})} x^{-\frac{n-p}{2}} y^{\frac{n-p}{2}-1}(1-\frac{y}{x})^{\frac{n(p-1)}{2}-1}\, dy \tag{4.22}$$
$$\times\, \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv\, dF(r).$$

The density of T^2 is

$$f_r(x) = \int_0^\infty \int_0^1 \int_0^x \Pr\left\{\mathbf{Y}'(r, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(r, v^{\frac{1}{2}}, \mathbf{u}^*) \leq \frac{y(1-v)}{n-1}\right\}$$
$$\times\, \frac{\Gamma(\frac{(n-1)p}{2})}{\Gamma(\frac{n-p}{2})\Gamma(\frac{n(p-1)}{2})} \left[-\frac{n-p}{2} x^{-\frac{n-p}{2}-1} y^{\frac{n-p}{2}-1}(1-\frac{y}{x})^{\frac{n(p-1)}{2}-1}\right.$$
$$\left. +(\frac{n(p-1)}{2}-1)x^{-\frac{n-p}{2}-2} y^{\frac{n-p}{2}}(1-\frac{y}{x})^{\frac{n(p-1)}{2}-2}\right] dy\, \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv\, dF(r)$$

$$= \frac{1}{x}\int_0^{\infty}\int_0^1\int_0^1 \Pr\left\{\mathbf{Y}'(r, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(r, v^{\frac{1}{2}}, \mathbf{u}^*) \leq \frac{x(1-v)w}{n-1}\right\}$$

$$\times \frac{n-p}{2}[\beta(w; \frac{n-p}{2}+1, \frac{n(p-1)}{2}-1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2})]\, dw \tag{4.23}$$

$$\times\, \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv\, dF(r).$$

We have the following theorem

Theorem 4.1.

If $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$, $n \geq p > 1$, and the density of $\mathbf{X}$ exists, then the density of T^2 has the form (4.23).

Next we want to attack the problem of LMPI by using Theorem 4.1. First let us define LMPI.

Definition 4.1.

A level α test ψ_0 is said to be locally most powerful if given any other level α test ψ, there exists δ such that the power of ψ_0 is bigger than the power of ψ (i.e. $P_{\psi_0}(\theta) \geq P_{\psi}(\theta)$ for all θ) with $0 < d(\theta) < \delta$, where $d(\theta)$ is a measure of the distance of θ from H_0.

Definition 4.2.

A level α test ψ_0 is said to be locally most powerful invariant (LMPI) if the test is locally most powerful among the family of all invariant tests.

Theorem 4.2.

Assume $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$, $n \geq p > 1$, and that the density of $\mathbf{X}$ exists. For testing $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$, the T^2-test is LMPI.

Proof.

Since, from Theorem 4.1, the non-null density of T^2 has the form (4.23),

$$\frac{\partial f_\tau(x)}{\partial \tau} = \frac{1}{x}\int_0^1\int_0^1 \frac{\partial}{\partial \tau}\Pr\left\{\mathbf{Y}'(R, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(R, v^{\frac{1}{2}}, \mathbf{u}^*) \le \frac{x(1-v)w}{n-1}\right\}$$

$$\times \frac{n-p}{2}[\beta(w; \frac{n-p}{2}+1, \frac{n(p-1)}{2}-1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2})]\, dw \tag{4.24}$$

$$\times\, \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv.$$

But

$$\mathbf{Y}'(R, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(R, v^{\frac{1}{2}}, \mathbf{u}^*) = \frac{\tau^2}{R^2} + 2\frac{(nv)^{\frac{1}{2}}}{R}(\mu-\mu_0)'\mathbf{D}^{-1}\mathbf{u}^* + v, \tag{4.25}$$

where $\tau^2 = n(\mu-\mu_0)'\Sigma^{-1}(\mu-\mu_0)$. There exists a $p \times p$ orthogonal matrix $\mathbf{Q}_1$ with first row $\frac{n^{\frac{1}{2}}(\mu-\mu_0)'\mathbf{D}^{-1}}{\tau}$ such that

$$\mathbf{Q}_1\mathbf{u}^* \stackrel{d}{=} \begin{pmatrix} BZ_1^{\frac{1}{2}} \\ \mathbf{Z}_2 \end{pmatrix} \tag{4.26}$$

$$\stackrel{d}{=} \mathbf{u}^{(p)},$$

where $\Pr\{B=1\} = \frac{1}{2} = \Pr\{B=-1\}$, $Z_1 \sim \text{Beta}(\frac{1}{2}, \frac{p-1}{2})$, and B, Z_1 are independent. So

$$\mathbf{Y}'(R, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(R, v^{\frac{1}{2}}, \mathbf{u}^*) \stackrel{d}{=} \frac{\tau^2}{R^2} + 2\frac{\tau}{R}(vZ_1)^{\frac{1}{2}}B + v, \tag{4.27}$$

and

$$\Pr\left\{\mathbf{Y}'(R, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(R, v^{\frac{1}{2}}, \mathbf{u}^*) \le \frac{x(1-v)w}{n-1}\right\}$$

$$= \Pr\left\{\frac{r^2}{R^2} + 2\frac{r}{R}(vZ_1)^{\frac{1}{2}}B + v \le \frac{x(1-v)w}{n-1}\right\}$$

$$= \int_0^1 \frac{1}{2}\beta(z_1; \frac{1}{2}, \frac{p-1}{2}) \left[\Pr\left\{\frac{r^2}{R^2} + 2\frac{r}{R}(vz_1)^{\frac{1}{2}} + v \le \frac{x(1-v)w}{n-1}\right\}\right. \tag{4.28}$$

$$\left. + \Pr\left\{\frac{r^2}{R^2} - 2\frac{r}{R}(vz_1)^{\frac{1}{2}} + v \le \frac{x(1-v)w}{n-1}\right\}\right] dz_1,$$

where $0 \le v, w \le 1$.

Since, under $0 \le v, w, z_1 \le 1$ and $\Pr\{R > 0\} = 1$,

$$\frac{r^2}{R^2} + 2\frac{r}{R}(vz_1)^{\frac{1}{2}} + v \le \frac{x(1-v)w}{n-1}$$

$$\Leftrightarrow (\frac{r}{R} + (vz_1)^{\frac{1}{2}})^2 + v(1-z_1) \le \frac{x(1-v)w}{n-1}$$

$$\Leftrightarrow 0 \le (\frac{r}{R} + (vz_1)^{\frac{1}{2}})^2 \le \frac{x(1-v)w}{n-1} - v(1-z_1)$$

$$\Leftrightarrow 0 \le (\frac{r}{R} + (vz_1)^{\frac{1}{2}}) \le [\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} \quad \text{and} \tag{4.29}$$
$$\frac{x(1-v)w}{n-1} \ge v(1-z_1)$$

$$\Leftrightarrow 0 \le \frac{r}{R} \le [\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} - (vz_1)^{\frac{1}{2}} \quad \text{and}$$
$$\frac{x(1-v)w}{n-1} \ge v(1-z_1),$$

and

$$\frac{r^2}{R^2} - 2\frac{r}{R}(vz_1)^{\frac{1}{2}} + v \le \frac{x(1-v)w}{n-1}$$

$$\Leftrightarrow (\frac{r}{R} - (vz_1)^{\frac{1}{2}})^2 + v(1-z_1) \le \frac{x(1-v)w}{n-1}$$

$$\Leftrightarrow -[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} \le \frac{r}{R} - (vz_1)^{\frac{1}{2}} \le [\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}}, \quad \frac{x(1-v)w}{n-1} \ge v(1-z_1) \tag{4.30}$$

$$\Leftrightarrow (vz_1)^{\frac{1}{2}} - [\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} \le \frac{r}{R} \le (vz_1)^{\frac{1}{2}} + [\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}}, \quad \frac{x(1-v)w}{n-1} \ge v(1-z_1),$$

$$\begin{aligned}
&\Pr\left\{\mathbf{Y}'(R, v^{\frac{1}{2}}, \mathbf{u}^*)\mathbf{Y}(R, v^{\frac{1}{2}}, \mathbf{u}^*) \le \frac{x(1-v)w}{n-1}\right\} \\
&= \int_0^1 \left[\frac{1}{2} I_{A \bigcap F} \left(\Pr\left\{R \ge \frac{r}{[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} - (vz_1)^{\frac{1}{2}}}\right\}\right.\right. \\
&\quad \left. + \Pr\left\{R \ge \frac{r}{[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} + (vz_1)^{\frac{1}{2}}}\right\}\right) + \frac{1}{2} I_{A \bigcap F^c} \\
&\quad \left. \times \Pr\left\{\frac{r}{[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} + (vz_1)^{\frac{1}{2}}} \le R \le \frac{r}{(vz_1)^{\frac{1}{2}} - [\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}}}\right\}\right] \\
&\quad \times \beta(z_1; \frac{1}{2}, \frac{p-1}{2})\, dz_1,
\end{aligned} \tag{4.31}$$

where $A = \left\{\frac{x(1-v)w}{n-1} > v(1-z_1)\right\}$, $F = \left\{[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} > (vz_1)^{\frac{1}{2}}\right\}$,

and $0 \le v, w \le 1$.

We have, from (4.23),

$$\frac{\partial f_r(x)}{\partial r}\Big|_{r=0}$$

$$= \frac{n-p}{2}\frac{1}{x}\int_0^1\int_0^1\int_0^1 \Bigg[\frac{1}{2}I_{A\bigcap F}\left(\frac{-1}{[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} - (vz_1)^{\frac{1}{2}}}f(0)\right.$$

$$\left. - \frac{1}{[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} + (vz_1)^{\frac{1}{2}}}f(0)\right)$$

$$+ \frac{1}{2}I_{A\bigcap F^c}\left(\frac{1}{(vz_1)^{\frac{1}{2}} - [\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}}}f(0)\right.$$

$$\left. - \frac{1}{[\frac{x(1-v)w}{n-1} - v(1-z_1)]^{\frac{1}{2}} + (vz_1)^{\frac{1}{2}}}f(0)\right)\Bigg]$$

$$\times \beta(z_1; \frac{1}{2}, \frac{p-1}{2})\, dz_1[\beta(w; \frac{n-p}{2}+1, \frac{n(p-1)}{2}-1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2})]\, dw$$

$$\times \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv$$

$$= \frac{n-p}{2}\frac{1}{x}f(0)\int_0^1\int_0^1\int_0^1 I_A \frac{-(\frac{x(1-v)w}{n-1} - v(1-z_1))^{\frac{1}{2}}}{\frac{x(1-v)w}{n-1} - v}$$

$$\times \beta(z_1; \frac{1}{2}, \frac{p-1}{2})\, dz_1[\beta(w; \frac{n-p}{2}+1, \frac{n(p-1)}{2}-1) - \beta(w; \frac{n-p}{2}, \frac{n(p-1)}{2})]\, dw$$

$$\times \beta(v; \frac{p}{2}, \frac{(n-1)p}{2})\, dv, \tag{4.32}$$

where $f(\cdot)$ is the density function of R.

Thus $\frac{\frac{\partial f_\tau(x)}{\partial \tau}|_{\tau=0}}{f(0) f_0(x)}$ does not depend on the density of R. Since $f(0)$ is constant and ,under multivariate normality, the T^2-test is LMPI, the T^2-test is LMPI in multivariate elliptically contoured case. ∎

Kariya (1981a, 1981b) has derived that T^2-test is UMPI under a subfamily of multivariate elliptically contoured distribution (g (see (3.2)) is nonincreasing convex function.) For the more general case as the one in Theorem 4.2, the property of UMPI is still an open problem.

Chapter 5

TESTING VARIOUS OTHER HYPOTHESIS

In this chapter, invariant tests for testing various hypotheses are studied. There are five sections in this chapter. In each section, the invariant tests and their properties in different hypothesis testing are derived. The order of these problems are test for lack of correlations among the sets, test of equality of covariance matrices and mean vectors, test for known correlation coefficient, test for known partial correlation coefficient, and test for zero multiple correlation coefficient.

5.1. Criteria for testing lack of correlations among the sets.

Consider $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$, and $n \geq p$. If the density of $\mathbf{X}$ exists, then the density of $\mathbf{X}$ has the form of

$$|\Sigma|^{-\frac{1}{2}n}\, g(\operatorname{tr}\Sigma^{-1}(\mathbf{X}-\mathbf{1}_n\mu')'(\mathbf{X}-\mathbf{1}_n\mu')), \tag{5.1}$$

where $\mathbf{1}_n' = (1, \cdots, 1)$.

Suppose we partition

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & \Sigma_{1q} \\ \vdots & & \vdots \\ \Sigma_{q1} & \cdots & \Sigma_{qq} \end{pmatrix}, \tag{5.2}$$

such that Σ_{ij}, $i = 1, \ldots, q$, $j = 1, \ldots, q$ is $p_i \times p_j$ matrix and $\sum_{i=1}^{q} p_i = p$. We wish to test $H_0 : \Sigma_{ij} = 0$, $i \neq j$ against the general alternative $H_1 : \Sigma_{ij} \neq 0$ for some $i \neq j$. First we have to find the sufficient statistics. Since

$$
\begin{aligned}
&(\mathbf{X}-\mathbf{1}_n\mu')'(\mathbf{X}-\mathbf{1}_n\mu') \\
&= (\mathbf{X}-\mathbf{1}_n\bar{\mathbf{x}}')'(\mathbf{X}-\mathbf{1}_n\bar{\mathbf{x}}') + (\mathbf{1}_n(\bar{\mathbf{x}}-\mu)')'(\mathbf{1}_n(\bar{\mathbf{x}}-\mu)') \\
&= \mathbf{A} + n(\bar{\mathbf{x}}-\mu)(\bar{\mathbf{x}}-\mu)',
\end{aligned} \tag{5.3}
$$

the density of $\mathbf{X}$ has the form of

$$
|\Sigma|^{-\frac{1}{2}n}\, g(\operatorname{tr}\Sigma^{-1}[\mathbf{A}+n(\bar{\mathbf{x}}-\mu)(\bar{\mathbf{x}}-\mu)']), \tag{5.4}
$$

where $\bar{\mathbf{x}} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n$ and $\mathbf{A} = (\mathbf{X}-\mathbf{1}_n\bar{\mathbf{x}}')'(\mathbf{X}-\mathbf{1}_n\bar{\mathbf{x}}')$. By the Factorization theorem, $\mathbf{A}$ and $\bar{\mathbf{x}}$ form a sufficient set of statistics for Σ and μ. This problem remains invariant under

$$
\begin{aligned}
&(1)\mathbf{X} \mapsto \mathbf{X} + \mathbf{1}_n\mathbf{d}', \\
&(2)\mathbf{X} \mapsto \mathbf{X}\mathbf{C},
\end{aligned} \tag{5.5}
$$

where $\mathbf{C}$ is a matrix with nonsingular diagonal blocks of $\mathbf{C}_i$ with order p_i, $i = 1, \ldots, q$, and off-diagonal blocks of 0's.

Under group (1),

$$
\mathbf{A} \mapsto \mathbf{A}, \qquad \bar{\mathbf{x}} \mapsto \bar{\mathbf{x}} + \mathbf{d}. \tag{5.6}
$$

So the only invariants of the sufficient statistics are $\mathbf{A}$ (by letting $\mathbf{d} = -\bar{\mathbf{x}}$.) Under group (2),

$$
\mathbf{A} \mapsto \begin{pmatrix} \mathbf{C}_1'\mathbf{A}_{11}\mathbf{C}_1 & \cdots & \mathbf{C}_1'\mathbf{A}_{1q}\mathbf{C}_q \\ \vdots & & \vdots \\ \mathbf{C}_q'\mathbf{A}_{q1}\mathbf{C}_1 & \cdots & \mathbf{C}_q'\mathbf{A}_{qq}\mathbf{C}_q \end{pmatrix}, \tag{5.7}
$$

where $\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1q} \\ \vdots & & \vdots \\ \mathbf{A}_{q1} & \cdots & \mathbf{A}_{qq} \end{pmatrix}$ and $\mathbf{A}_{ij}$ is $p_i \times p_j$ matrix. Let $\mathbf{A}_0$ be a matrix with diagonal blocks of $\mathbf{A}_{ii}$, $i = 1, \ldots, q$, and off-diagonal blocks of 0's, $\mathbf{B}$ be a matrix with diagonal blocks

of $\mathbf{B}_i$, $i = 1, \ldots, q$, and off-diagonal blocks of 0's such that $\mathbf{B}_i'\mathbf{A}_{ii}\mathbf{B}_i = \mathbf{I}$, $i = 1, \ldots, q$, then $\mathbf{B}'\mathbf{A}_0\mathbf{B} = \mathbf{I}$. The invariants of the sufficient statistics are

$$\mathbf{B}'\mathbf{A}\mathbf{B} = \begin{pmatrix} \mathbf{I} & \mathbf{B}_1'\mathbf{A}_{12}\mathbf{B}_2 & \cdots & \mathbf{B}_1'\mathbf{A}_{1q}\mathbf{B}_q \\ \mathbf{B}_2'\mathbf{A}_{21}\mathbf{B}_1 & \mathbf{I} & \cdots & \mathbf{B}_2'\mathbf{A}_{2q}\mathbf{B}_q \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_q'\mathbf{A}_{q1}\mathbf{B}_1 & \mathbf{B}_q'\mathbf{A}_{q2}\mathbf{B}_2 & \cdots & \mathbf{I} \end{pmatrix}. \tag{5.8}$$

So $\mathbf{Q}'\mathbf{B}'(\mathbf{A} - \mathbf{A}_0)\mathbf{B}\mathbf{Q}$ is invariant with respect to group (2), where $\mathbf{Q}$ is a matrix with orthogonal diagonal blocks and off-diagonal blocks of 0's. Hence we conclude that the invariant tests for the multivariate elliptically contoured distributions are same as the ones in multivariate normal case.

Some invariant test criteria are :

(1) Likelihood ratio criterion :

$$\left|\mathbf{B}'(\mathbf{A} - \mathbf{A}_0)\mathbf{B} + \mathbf{I}\right|. \tag{5.9}$$

(2) Nagao's criterion :

$$\frac{1}{2}\,\mathrm{tr}[\mathbf{B}'(\mathbf{A} - \mathbf{A}_0)\mathbf{B}]^2. \tag{5.10}$$

Since the likelihood ratio criterion is admissible in the multivariate normal case, from Theorem 3.2 and Theorem 3.4, we know that the likelihood ratio criterion is also admissible in the multivariate elliptically contoured case.

Next we want to find the null and non-null distributions for the above criteria. From Corollary 2.1, we have the following corollary.

Corollary 5.1.

For both the null and non-null cases, the distribution of $\mathbf{B}'(\mathbf{A} - \mathbf{A}_0)\mathbf{B}$ *does not depend on any particular underlying elliptically contoured distribution, i.e. the null and non-null distributions are same as the ones in the multivariate normal case.*

From the above corollary, we know that the distributions of likelihood ratio criterion and Nagao's criterion are the same as the ones in multivariate normal case. So their asymptotic distributions are the same as the ones in multivariate normal case.

5.2. Testing of equality of covariance matrices and mean vectors.

Consider

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}'_{11} \\ \vdots \\ \mathbf{x}'_{1n_1} \\ \vdots \\ \mathbf{x}'_{q1} \\ \vdots \\ \mathbf{x}'_{qn_q} \end{pmatrix} \sim MEC_{n\times p}(\mathbf{M}; \Sigma_1, \ \dots \ , \Sigma_1, \Sigma_2, \ \dots \ , \Sigma_2, \ \dots \ , \Sigma_q, \ \dots \ , \Sigma_q; \phi)$$

with n_i Σ_i's, $\Sigma_i > 0$, $n_i > p$, $i = 1, \dots, q$, $\sum_{i=1}^q n_i = n$, $\mathbf{x}_{ij}$; $j = 1, \dots, n_i$, $i = 1, \dots, q$, are $p \times 1$ vectors, the rows of $\mathbf{M}$ are n_i (μ'_i)'s, $i = 1, \dots, q$, and the density of $\mathbf{X}$ exists. Then the density of $\mathbf{X}$ has the following form

$$(\prod_{i=1}^q |\Sigma_i|^{-\frac{1}{2}n_i}) g(\sum_{i=1}^q \operatorname{tr} \Sigma_i^{-1} \mathbf{G}_i), \tag{5.11}$$

where

$$\begin{aligned} \mathbf{G}_i &= \sum_{j=1}^{n_i} (\mathbf{x}_{ij} - \mu_i)(\mathbf{x}_{ij} - \mu_i)' \\ &= \mathbf{A}_i + n_i(\bar{\mathbf{x}}_i - \mu_i)(\bar{\mathbf{x}}_i - \mu_i)', \end{aligned}$$

$$\mathbf{A}_i = \sum_{j=1}^{n_i} (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)',$$

and

$$\bar{\mathbf{x}}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{x}_{ij}.$$

By the Factorization theorem, $\mathbf{A}_i$, $\bar{\mathbf{x}}_i$, $i = 1,\ldots,q$, form a sufficient set of statistics for Σ_i, μ_i, $i = 1\ldots,q$. Now we wish to test $H_0 : \mu_1 = \cdots = \mu_q$ and $\Sigma_1 = \cdots = \Sigma_q$ against the general alternative $H_1 : \mu_i \neq \mu_j$ for some $i \neq j$ or $\Sigma_i \neq \Sigma_j$ for some $i \neq j$.

The problem remains invariant under

$$\begin{aligned}
&(1)\mathbf{x}_{ij} \mapsto \mathbf{x}_{ij} + \mathbf{d}, \qquad \forall i,j,\\
&(2)\mathbf{D} \mapsto \mathbf{D}\mathbf{C}^{*-1}, \qquad \mathbf{B} \mapsto \mathbf{C}^*\mathbf{B},\\
&(3)\mathbf{x}_{ij} \mapsto \mathbf{C}\mathbf{x}_{ij}, \qquad \forall i,j,
\end{aligned} \tag{5.12}$$

where $\mathbf{C}^*$ and $\mathbf{C}$ are nonsingular, $\mathbf{D} = (\mu_1 - \mu_q, \cdots, \mu_{q-1} - \mu_q)$, and $\mathbf{B}$ is a $(q-1)\times n$ matrix with (i, j)th element 1 when $\sum_{k=1}^{i-1} n_k < j \leq \sum_{k=1}^{i} n_k$, 0 otherwise.

Under group (1),

$$\mathbf{A}_i \mapsto \mathbf{A}_i, \qquad \bar{\mathbf{x}}_i \mapsto \bar{\mathbf{x}}_i + \mathbf{d}, \qquad \forall i. \tag{5.13}$$

Hence the invariants of the sufficient statistics are $(\mathbf{A}_1,\cdots,\mathbf{A}_q)$ and $(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_q, \cdots, \bar{\mathbf{x}}_{q-1} - \bar{\mathbf{x}}_q)$. Under group (2), as in MANOVA, the invariants of the sufficient statistics are $(\mathbf{A}_1,\cdots,\mathbf{A}_q)$ and

$$\mathbf{H} = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_q, \cdots, \bar{\mathbf{x}}_{q-1} - \bar{\mathbf{x}}_q)\mathbf{E}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_q, \cdots, \bar{\mathbf{x}}_{q-1} - \bar{\mathbf{x}}_q)',$$

where $\mathbf{E} = (e_{ij})$ is a $(q-1)\times(q-1)$ matrix with the element $e_{ij} = \delta_{ij}n_i - \frac{1}{n}n_i n_j$. Since

$$\begin{aligned}
\mathbf{A} &= \sum_{i=1}^{q}\sum_{j=1}^{n_i}(\mathbf{x}_{ij} - \bar{\mathbf{x}})(\mathbf{x}_{ij} - \bar{\mathbf{x}})'\\
&= \mathbf{H} + \sum_{i=1}^{q}\mathbf{A}_i,
\end{aligned} \tag{5.14}$$

where $\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{q}\sum_{j=1}^{n_i}\mathbf{x}_{ij}$, the invariants of the sufficient statistics are $\mathbf{A}$, $\mathbf{A}_1, \cdots, \mathbf{A}_q$. Under group (3),

$$\mathbf{A} \mapsto \mathbf{CAC}', \qquad \mathbf{A}_i \mapsto \mathbf{CA}_i\mathbf{C}', \qquad \forall i, \tag{5.15}$$

where $\mathbf{C}$ is nonsingular.

Let $\mathbf{CAC}' = \mathbf{I}$, then $\mathbf{QCA}_i\mathbf{C}'\mathbf{Q}'$, $i = 1, \ldots, q$, is invariant with respect to group (3), where $\mathbf{Q}$ is an $p \times p$ orthogonal matrix. Therefore we conclude that the invariant tests in multivariate elliptically contoured distribution are same as the ones in multivariate normal case.

Some invariant test criteria are :

(1) Likelihood ratio criterion :

$$\frac{\prod_{i=1}^{q}|\mathbf{A}_i|^{\frac{1}{2}n_i}}{|\mathbf{A}|^{\frac{1}{2}n}}. \tag{5.16}$$

(2) Bartlett modified likelihood ratio criterion :

$$\frac{\prod_{i=1}^{q}|\mathbf{A}_i|^{\frac{1}{2}(n_i-1)}}{|\mathbf{A}|^{\frac{1}{2}(n-q)}}. \tag{5.17}$$

For dealing with the null distribution we have the following corollary from Theorem 2.1.

Corollary 5.2.

Under $\mu_1 = \cdots = \mu_q$, the joint distribution of $\mathbf{CA}_i\mathbf{C}'$, $i = 1, \ldots, q$, does not depend on any particular underlying elliptically contoured distribution, i.e. it is same as the one in the multivariate normal case.

From the above corollary, the null distribution of the invariant test criterion in the multivariate elliptically contoured distribution is the same as the one in the multivariate normal case.

5.3. Test of Hypothesis about correlation coefficient.

Assume $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$ and $n \geq p$, and the density of $\mathbf{X}$ exists, where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, $\Sigma_{11} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ is a 2×2 matrix. We wish to test $H_0 : \rho = \rho_0$, where $|\rho_0| < 1$ is a given number.

Let $\mathbf{B} = \begin{pmatrix} \mathbf{I}_{2\times 2} \\ \mathbf{0} \end{pmatrix}$ be a $p \times 2$ matrix, then from Corollary 1 of Anderson and Fang (1982b), $\mathbf{XB} \sim LEC_{n\times 2}(\mu\mathbf{B}, \Sigma_{11}, \phi^*)$ with $\phi^* \in \Phi_{2n} \leftrightarrow R^* = Rb_{n,\frac{1}{2}n(p-2)}$, $b^2_{n,\frac{1}{2}n(p-2)} \sim$ Beta$(n, \frac{1}{2}n(p-2))$. So ,without loss of generality, let $p = 2$. i.e. $\mathbf{X}_{n\times 2} \sim LEC_{n\times 2}(\mu, \Sigma, \phi)$, with $\Sigma > 0$, and $n \geq 2$, and let the density of $\mathbf{X}$ exist, where $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. The density of $\mathbf{X}$ has the form of

$$|\Sigma|^{-\frac{1}{2}n} g(\mathrm{tr}\,\Sigma^{-1}\mathbf{G}), \tag{5.18}$$

where

$$\begin{aligned} \mathbf{G} &= (\mathbf{X} - \mathbf{1}_n\mu')'(\mathbf{X} - \mathbf{1}_n\mu') \\ &= \mathbf{A} + n(\bar{\mathbf{x}} - \mu)(\bar{\mathbf{x}} - \mu)', \end{aligned}$$

$$\bar{\mathbf{x}} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n,$$

and

$$\begin{aligned} \mathbf{A} &= (\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}') \\ &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \end{aligned} \tag{5.19}$$

By the Factorization theorem, $\mathbf{A}$ and $\bar{\mathbf{x}}$ form a sufficient set of statistics for Σ and μ. The problem remains invariant under

$$(1)\mathbf{X} \mapsto \mathbf{X} + \mathbf{1}_n\mathbf{d}',$$

$$(2)\mathbf{X} \mapsto \mathbf{X}\mathbf{C}, \tag{5.20}$$

where $\mathbf{C} = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}$, and $c_1, c_2 > 0$.

Under group (1),

$$\mathbf{A} \mapsto \mathbf{A}, \qquad \bar{\mathbf{x}} \mapsto \bar{\mathbf{x}} + \mathbf{d}, \tag{5.21}$$

and the only invariants of the sufficient statistics are $\mathbf{A}$ (by letting $\mathbf{d} = -\bar{\mathbf{x}}$.) Under group (2),

$$\mathbf{A} \mapsto \begin{pmatrix} c_1^2 a_{11} & c_1c_2a_{12} \\ c_1c_2a_{21} & c_2^2 a_{22} \end{pmatrix}. \tag{5.22}$$

The only invariant (i.e. maximal invariant) of the sufficient statistics is $r_{12} = \frac{a_{12}}{a_{11}^{\frac{1}{2}}a_{22}^{\frac{1}{2}}}$ (by letting $c_1 = a_{11}^{-\frac{1}{2}}$, $c_2 = a_{22}^{-\frac{1}{2}}$.) We conclude that the invariant tests in multivariate elliptically contoured distribution are same as the ones in multivariate normal case.

For finding the null and non-null distributions of r_{12}, from Corollary 2.1, we have the following corollary.

Corollary 5.3.

For both the null and non-null cases, the distribution of r_{12} does not depend on any particular underlying multivariate elliptically contoured distribution.

From the above corollary, all the optimum properties of r_{12} for multivariate normal case is fulfilled for multivariate elliptically contoured case, such as : asymptotic normality, and "Fisher's z" test. From Corollary 2.2, r_{12} is UMPI against alternatives $\rho > \rho_0$.

5.4. Test of hypothesis about partial correlation coefficient.

Assume $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$ and $n \geq p > q$, and the density of $\mathbf{X}$ exists, where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, and Σ_{11} is a $q \times q$ matrix. We wish to test H_0 : $\rho_{12\cdot q+1,\cdots,p} = \rho_0$, where $|\rho_0| < 1$ is a given number. Without loss of generality, assume that $q = 2$ (by the same technique as in section 5.3.)

Let $\mathbf{X} = (\mathbf{X}_1, \quad \mathbf{X}_2)$, where $\mathbf{X}_1 = (\mathbf{x}_1, \quad \mathbf{x}_2)$ is $n \times 2$, $\mathbf{X}_2$ is $n \times (p-2)$, and $\mathbf{x}_1$, $\mathbf{x}_2$ are $n \times 1$ vectors. The density of $\mathbf{X}$ has the form of

$$|\Sigma|^{-\frac{1}{2}n} g(\operatorname{tr} \Sigma^{-1}\mathbf{G}), \tag{5.23}$$

where

$$\begin{aligned} \mathbf{G} &= (\mathbf{X} - \mathbf{1}_n\mu')'(\mathbf{X} - \mathbf{1}_n\mu') \\ &= \mathbf{A} + n(\bar{\mathbf{x}} - \mu)(\bar{\mathbf{x}} - \mu)', \end{aligned}$$

$$\bar{\mathbf{x}} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n,$$

$$\begin{aligned} \mathbf{A} &= (\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}') \\ &= \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \end{aligned} \tag{5.24}$$

and $\mathbf{A}_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is a 2×2 matrix.

By the Factorization theorem, $\mathbf{A}$ and $\bar{\mathbf{x}}$ form a sufficient set of statistics for Σ and μ. The problem remains invariant under

$$
\begin{aligned}
&(1)\mathbf{X} \mapsto \mathbf{X} + \mathbf{1}_n\mathbf{d}',\\
&(2)\mathbf{X} \mapsto \mathbf{X}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix},\\
&(3)\mathbf{X} \mapsto \mathbf{X}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B} & \mathbf{I} \end{pmatrix},\\
&(4)\mathbf{X}_1 \mapsto \mathbf{X}_1\begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix},
\end{aligned}
\tag{5.25}
$$

where $\mathbf{C}$ is a $(p-2)\times(p-2)$ non-singular matrix, $\mathbf{B}$ is a $(p-2)\times 2$ matrix, and $b_1, b_2 > 0$.

Under group (1),

$$\mathbf{A} \mapsto \mathbf{A}, \qquad \bar{\mathbf{x}} \mapsto \bar{\mathbf{x}} + \mathbf{d}. \tag{5.26}$$

So the only invariants of the sufficient statistics are $\mathbf{A}$ (by letting $\mathbf{d} = -\bar{\mathbf{x}}$). Under group (2),

$$\mathbf{A} \mapsto \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12}\mathbf{C} \\ \mathbf{C}'\mathbf{A}_{21} & \mathbf{C}'\mathbf{A}_{22}\mathbf{C} \end{pmatrix}. \tag{5.27}$$

Hence the only invariants of the sufficient statistics are $\mathbf{A}_{11}$, $\mathbf{A}_{12}\mathbf{C}$ (by letting $\mathbf{C}'\mathbf{A}_{21}\mathbf{C} = \mathbf{I}$.) Under group (3),

$$
\begin{aligned}
&\mathbf{A}_{11} \mapsto \mathbf{A}_{11} + \mathbf{B}'\mathbf{C}'\mathbf{A}_{21} + \mathbf{A}_{12}\mathbf{C}\mathbf{B} + \mathbf{B}'\mathbf{B},\\
&\mathbf{A}_{12}\mathbf{C} \mapsto \mathbf{A}_{12}\mathbf{C} + \mathbf{B}'.
\end{aligned}
\tag{5.28}
$$

The only invariants of the sufficient statistics are

$$\begin{aligned}\mathbf{A}_{11\cdot 2} &= \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \\ &= \begin{pmatrix} a_{11\cdot 3,\cdots,p} & a_{12\cdot 3,\cdots,p} \\ a_{21\cdot 3,\cdots,p} & a_{22\cdot 3,\cdots,p} \end{pmatrix}\end{aligned}$$

(by letting $\mathbf{B}' = -\mathbf{A}_{12}\mathbf{C}$.) Under group (4),

$$\mathbf{A}_{11\cdot 2} \mapsto \begin{pmatrix} b_1^2 a_{11\cdot 3,\cdots,p} & b_1 b_2 a_{12\cdot 3,\cdots,p} \\ b_1 b_2 a_{21\cdot 3,\cdots,p} & b_2^2 a_{22\cdot 3,\cdots,p} \end{pmatrix}. \tag{5.29}$$

Thus the only invariant (i.e. maximal invariant) of the sufficient statistics is $r_{12\cdot 3,\cdots,p} = \frac{a_{12\cdot 3,\cdots,p}}{a_{11\cdot 3,\cdots,p}^{\frac{1}{2}} a_{22\cdot 3,\cdots,p}^{\frac{1}{2}}}$ (by letting $b_i = a_{ii\cdot 3,\cdots,p}^{-\frac{1}{2}}$, $i = 1,2$.) We conclude that the invariant test in multivariate elliptically contoured distribution is same as the one in multivariate normal case.

Next we want to find the null and non-null distributions of $r_{12\cdot 3,\cdots,p}$ for multivariate elliptically contoured distribution.

Corollary 5.4.

For both the null and non-null cases, the distribution of $r_{12\cdot 3,\cdots,p}$ does not depend on the underlying multivariate elliptically contoured distribution.

Proof.

Clear from Corollary 2.1. ∎

From the above corollary and Theorem 4.3.5 due to Anderson (1984), the distribution of a sample partial correlation coefficient, $r_{12\cdot 3,\cdots,p}$, based on $\mathbf{X}_{n\times p}$ with population partial correlation coefficient $\rho_{12\cdot 3,\cdots,p}$ equal to a certain value, ρ, is the same as the distribution of an ordinary correlation coefficient based on $\mathbf{X}_{[n-(p-q)]\times p}$ with corresponding population correlation of ρ. Also the asymptotic normality and "Fisher's z" test are fulfilled for the multivariate elliptically contoured case. Obviously, $r_{12\cdot 3,\cdots,p}$ is UMPI against alternatives $\rho_{12\cdot 3,\cdots,p} > \rho_0$.

5.5. Test of hypothesis about the multiple correlation coefficient.

Assume $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$, with $\Sigma > 0$ and $n \geq p$, and the density of $\mathbf{X}$ exists, where $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, and Σ_{11} is a $q \times q$ matrix. We wish to test $H_0 : \bar{R}_{1\cdot q+1,\cdots,p} = 0$, against the general alternative $H_1 : \bar{R}_{1\cdot q+1,\cdots,p} > 0$. So ,without loss of generality, assume that $q = 1$ (by the same technique as the one in section 5.3.)

Let $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma'_{(1)} \\ \sigma_{(1)} & \Sigma_{22} \end{pmatrix}$, then $\bar{R}_{1\cdot 2,\cdots,p} = \dfrac{(\sigma'_{(1)}\Sigma_{22}^{-1}\sigma_{(1)})^{\frac{1}{2}}}{\sigma_{11}^{\frac{1}{2}}} = 0$ iff $\sigma_{(1)} = 0$. The test is same as $H_0 : \sigma_{(1)} = 0$ against $H_1 : \sigma_{(1)} \neq 0$. This is a special case of testing for lack of correlations among the sets. Then, by the result in section 5.1, $\mathbf{B}'_2\mathbf{a}_{(1)}b_1$ are the invariants of the sufficient statistics, where $\mathbf{B}'_2\mathbf{A}_{22}\mathbf{B}_2 = \mathbf{I}$, $b_1^2 a_{11} = 1$, and $\mathbf{A} = \begin{pmatrix} a_{11} & \mathbf{a}'_{(1)} \\ \mathbf{a}_{(1)} & \mathbf{A}_{22} \end{pmatrix}$. Since the problem remains invariant under the group $\{\mathbf{Q} \mid \mathbf{X} \mapsto \mathbf{XQ}$, where $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{Q}_1 \end{pmatrix}$ and $\mathbf{Q}_1$ is $(p-1) \times (p-1)$ orthogonal $\}$. Under this group,

$$\mathbf{B}_2{}'\mathbf{a}_{(1)}b_1 \mapsto \mathbf{Q}'_1\mathbf{B}_2{}'\mathbf{a}_{(1)}b_1. \tag{5.30}$$

The only invariant (i.e. maximal invariant) of the sufficient statistics is $b_1(\mathbf{a}'_{(1)}\mathbf{B}_2\mathbf{B}'_2\mathbf{a}_{(1)})^{\frac{1}{2}} = \dfrac{(\mathbf{a}'_{(1)}\mathbf{A}_{22}^{-1}\mathbf{a}_{(1)})^{\frac{1}{2}}}{a_{11}^{\frac{1}{2}}} = \bar{R}$ (by letting the first column of $\mathbf{Q}_1$ is $\mathbf{B}'_2\mathbf{a}_{(1)}$.) Also ,by Corollary 5.1, we claim that the null and non-null distributions of $\bar{R}$ in multivariate elliptically contoured distribution are same as the ones in multivariate normal case. So $\bar{R}$ is UMPI. Also, the invariant test is same as the one in multivariate normal case.

Chapter 6

MAXIMUM LIKELIHOOD ESTIMATES AND LIKELIHOOD RATIO TESTS

In this chapter maximum likelihood estimates (MLE) and likelihood ratio tests (LRT) for the multivariate elliptically contoured distribution are derived. We first give the result for a special case, then some examples, and finally the general theorem.

6.1. The result for a special case.

Anderson and Fang (1982c) attacked the problems of MLE and LRT and gave some conditions for finding MLE. In this section, more precise and weaker conditions are given. The following lemma contains the general idea for constructing the weaker conditions in finding MLE.

Lemma 6.1.

Assume that $g(\cdot)$ is a continuous function such that $g(x_1^2+\cdots+x_N^2)$ is the density of $EC_N(0,\mathrm{I},\phi)$ and $E(R^2)<\infty$ where $R \leftrightarrow \phi \in \Phi_N$. Then the function

$$h(x) \equiv x^{\frac{N}{2}} g(x), \qquad x \geq 0, \tag{6.1}$$

has a maximum at some finite $x_0 > 0$.

Proof.

Since, from Cambanis, Huang and Simons (1981) (or Lemma 13.3.1 of Anderson

(1984)), the density function of R^2 is

$$
\begin{aligned}
f(x) &\equiv \frac{\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2})} x^{\frac{1}{2}N-1} g(x), \\
h(x) &= \frac{\Gamma(\frac{N}{2})}{\pi^{\frac{N}{2}}} x f(x).
\end{aligned}
\tag{6.2}
$$

From $E(R^2) < \infty$, $\lim_{x\mapsto\infty} h(x) = 0$. Furthermore $h(0) = 0$. (If not, there exists $\epsilon > 0$, $xf(x) > \epsilon\ \forall x > \delta$, so $1 \geq \int_0^\delta f(x)\,dx > \int_0^\delta \frac{\epsilon}{x}\,dx = \infty$ which is a contradiction.) Hence $h(x) \geq 0\ \forall x \geq 0$, and $h(\cdot)$ is continuous. The assertion of the lemma follows. ∎

The following theorem is the main theorem in this section and let us define the assumption for this theorem.

Assumption A.

$\mathbf{X} \sim MEC_{n\times p}(\mathbf{M}; \Sigma^{(1)}, \ldots, \Sigma^{(1)}, \ldots, \Sigma^{(q)}, \ldots, \Sigma^{(q)}; \phi)$ *with* n_i $\Sigma^{(i)}$*'s*, $n_i > p$, $i = 1, \ldots, q$, *and* $\sum_{i=1}^q n_i = n$. *The rows of* $\mathbf{M}$ *are* n_1 $\mu^{(1)}$*'s*, ..., n_q $\mu^{(q)}$*'s, successively, where* $\Sigma^{(i)} = \mathrm{diag}(\Sigma_{11}^{(i)}, \ldots, \Sigma_{kk}^{(i)})$, $1 \leq k \leq p$. *The density function of* $\mathbf{X}$ *is*

$$
\left[\prod_{i=1}^{q}\prod_{j=1}^{k}\left|\Sigma_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right] g\left(\sum_{i=1}^{q}\sum_{j=1}^{k} \mathrm{tr}\, \Sigma_{jj}^{(i)-1} \mathbf{G}_{jj}^{(i)}\right), \tag{6.3}
$$

where $\mathbf{G}_i = \sum_{j=\bar{n}_{i-1}+1}^{\bar{n}_i} (\mathbf{x}_{(j)} - \mu^{(i)})(\mathbf{x}_{(j)} - \mu^{(i)})' = \begin{pmatrix} \mathbf{G}_{11}^{(i)} & \cdots & \mathbf{G}_{ik}^{(i)} \\ \vdots & & \vdots \\ \mathbf{G}_{k1}^{(i)} & \cdots & \mathbf{G}_{kk}^{(i)} \end{pmatrix}$ *with* $\bar{n}_0 = 0$, $\bar{n}_i = n_1 + \cdots + n_i$.

Theorem 6.1.

Under the assumption A, if $f(\lambda) \equiv \lambda^{-\frac{1}{2}np} g(\frac{p}{\lambda})$ *attains its maximum at some finite*

$\lambda_{\max}(g)$, then the MLEs of $\mu^{(i)}$ and $\Sigma^{(i)}$ are

$$\begin{aligned} \hat{\mu}^{(i)} &= \bar{\mathbf{x}}^{(i)}, \\ \hat{\Sigma}^{(i)} &= \operatorname{diag}(\hat{\Sigma}_{11}^{(i)}, \ldots, \hat{\Sigma}_{kk}^{(i)}), \end{aligned} \tag{6.4}$$

$i = 1, \ldots, q$, where $\hat{\Sigma}_{jj}^{(i)} = \left(\frac{n}{n_i}\right) \lambda_{\max}(g) \mathbf{W}_{jj}^{(i)}$, $i = 1, \ldots, q$, $j = 1, \ldots, k$. $\mathbf{W}_{jj}^{(i)}$ is obtained from $\mathbf{G}_{jj}^{(i)}$ by substituting $\mu^{(i)}$ with $\bar{\mathbf{x}}^{(i)} = \frac{1}{n_i} \sum_{j=n_{i-1}+1}^{n_i} \mathbf{x}_{(j)}$ and the corresponding maximized likelihood is

$$\lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \prod_{i=1}^{q} \left[\left(\frac{n}{n_i}\right)^{-\frac{1}{2}n_i p} \prod_{j=1}^{k} \left|\mathbf{W}_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right]. \tag{6.5}$$

Proof.

$$\begin{aligned} &\max_{\mu^{(i)}, \Sigma^{(i)}} L(\mu^{(1)}, \ldots, \mu^{(q)}; \Sigma^{(1)}, \ldots, \Sigma^{(q)}) \\ &= \max_{\mu^{(i)}} \max_{\Sigma^{(i)}} L(\mu^{(1)}, \ldots, \mu^{(q)}; \Sigma^{(1)}, \ldots, \Sigma^{(q)}), \end{aligned} \tag{6.6}$$

where

$$\begin{aligned} L \equiv& L(\mu^{(1)}, \ldots, \mu^{(q)}; \Sigma^{(1)}, \ldots, \Sigma^{(q)}) \\ =& \left[\prod_{i=1}^{q} \prod_{j=1}^{k} \left|\Sigma_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right] g\left(\sum_{i=1}^{q} \sum_{j=1}^{k} \operatorname{tr} \Sigma_{jj}^{(i)^{-1}} \mathbf{G}_{jj}^{(i)}\right). \end{aligned} \tag{6.7}$$

First we want to maximize L with respect to $\Sigma^{(i)}$, $i = 1, \ldots, q$. Since $\mathbf{G}_{jj}^{(i)} > 0$, $i = 1, \ldots, q$, $j = 1, \ldots, k$, with probability one, there exist nonsingular matrices $\mathbf{C}_{ij}$ such that $\mathbf{C}_{ij}\mathbf{C}_{ij}' = \mathbf{G}_{jj}^{(i)}$ and orthogonal matrices Γ_{ij} such that

$$\begin{aligned} \Gamma_{ij}' \mathbf{C}_{ij}^{-1} \Sigma_{jj}^{(i)} \mathbf{C}_{ij}'^{-1} \Gamma_{ij} &= \Lambda_{ij} \\ &= \operatorname{diag}(\lambda_{j1}^{(i)}, \ldots, \lambda_{jp_j}^{(i)}), \end{aligned}$$

with positive $\lambda_{j\alpha}^{(i)}$, $\alpha = 1, \ldots, p_j$. Therefore

$$L = \left[\prod_{i=1}^{q}\prod_{j=1}^{k}\left(\left|G_{jj}^{(i)}\right|\prod_{\alpha=1}^{p_j}\lambda_{j\alpha}^{(i)}\right)^{-\frac{1}{2}n_i}\right] g\left(\sum_{i=1}^{q}\sum_{j=1}^{k}\sum_{\alpha=1}^{p_j}1/\lambda_{j\alpha}^{(i)}\right). \tag{6.8}$$

It is a symmetric function of $\lambda_{j\alpha}^{(i)}$ $(j = 1, \ldots, k,\ \alpha = 1, \ldots, p_j)$ for $i = 1, \ldots, q$. Hence $\lambda_{j\alpha}^{(i)} = \lambda^{(i)}$, $j = 1, \ldots, k$, $\alpha = 1, \ldots, p_j$, and

$$L = \left[\left(\prod_{i=1}^{q}\prod_{j=1}^{k}\left|G_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right)\prod_{i=1}^{q}\lambda^{(i)-\frac{1}{2}n_ip}\right] g\left(p\sum_{i=1}^{q}\frac{1}{\lambda^{(i)}}\right). \tag{6.9}$$

Let

$$\frac{1}{\lambda} = \sum_{i=1}^{q}\frac{1}{\lambda^{(i)}},$$

$$\beta_j = \frac{\lambda}{\lambda^{(j)}}, \qquad j = 1, \ldots, q-1. \tag{6.10}$$

So maximizing L with respect to $\Sigma^{(i)}$, $i = 1, \ldots, q$, is the same as maximizing L with respect to λ, β_j, $j = 1, \ldots, q-1$.

Now

$$\begin{aligned} L &= \left(\prod_{i=1}^{q}\prod_{j=1}^{k}\left|G_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right)\left(\prod_{i=1}^{q}\lambda^{(i)-\frac{1}{2}n_ip}\right) g\left(p\sum_{i=1}^{q}\frac{1}{\lambda^{(i)}}\right) \\ &= \left(\prod_{i=1}^{q}\prod_{j=1}^{k}\left|G_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right)\left[\left(\prod_{i=1}^{q-1}\beta_i^{\frac{1}{2}n_ip}\right)\left(1-\sum_{i=1}^{q-1}\beta_i\right)^{\frac{1}{2}n_qp}\right]\lambda^{-\frac{1}{2}np}g\left(\frac{p}{\lambda}\right). \end{aligned} \tag{6.11}$$

So

$$\begin{aligned} &\max_{\lambda,\beta_1,\ldots,\beta_{q-1}} L \\ &= \prod_{i=1}^{q}\prod_{j=1}^{k}\left|G_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\left[\max_{\beta_1,\ldots,\beta_{q-1}}\left(\prod_{i=1}^{q-1}\beta_i^{\frac{1}{2}n_ip}\right)\left(1-\sum_{i=1}^{q-1}\beta_i\right)^{\frac{1}{2}n_qp}\right]\left[\max_{\lambda}\lambda^{-\frac{1}{2}np}g\left(\frac{p}{\lambda}\right)\right]. \end{aligned} \tag{6.12}$$

L attains its maximum at $\lambda = \lambda_{\max}(g)$, $\beta_i = \frac{n_i}{n}$, $i = 1, \ldots, q-1$ and

$$L^* = \max_{\Sigma^{(i)}} L$$
$$= \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \prod_{i=1}^{q} \left[\left(\frac{n}{n_i}\right)^{-\frac{1}{2}n_i p} \prod_{j=1}^{k} \left|\mathbf{G}_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right]. \tag{6.13}$$

Since $\mathbf{G}_i = \mathbf{W}_i + n_i(\bar{\mathbf{x}}^{(i)} - \mu^{(i)})(\bar{\mathbf{x}}^{(i)} - \mu^{(i)})'$, maximizing L^* with respect to $\mu^{(i)}$, $i = 1, \ldots, q$, shows that L attains its maximum at $\mu^{(i)} = \bar{\mathbf{x}}^{(i)}$. This completes the proof. ∎

Combining the above lemma and theorem, we have the following corollary.

Corollary 6.1.

Under the assumption A, if $g(\cdot)$ is continuous and $E(R^2) < \infty$ where $R \leftrightarrow \phi \in \Phi_{np}$, then the MLEs of $\mu^{(i)}$ and $\Sigma^{(i)}$ are

$$\begin{aligned} \hat{\mu}^{(i)} &= \bar{\mathbf{x}}^{(i)}, \\ \hat{\Sigma}^{(i)} &= \operatorname{diag}(\hat{\Sigma}_{11}^{(i)}, \ldots, \hat{\Sigma}_{kk}^{(i)}), \end{aligned} \tag{6.14}$$

$i = 1, \ldots, q$, where $\hat{\Sigma}_{jj}^{(i)} = (\frac{n}{n_i})\lambda_{\max}(g)\mathbf{W}_{jj}^{(i)}$, $i = 1, \ldots, q$, $j = 1, \ldots, k$, and the corresponding maximum of the likelihood function is

$$\lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \prod_{i=1}^{q} \left[\left(\frac{n}{n_i}\right)^{-\frac{1}{2}n_i p} \prod_{j=1}^{k} \left|\mathbf{W}_{jj}^{(i)}\right|^{-\frac{1}{2}n_i}\right]. \tag{6.15}$$

Under some conditions, Theorem 6.1 and Corollary 6.1 show that the maximum of the likelihood function for the multivariate elliptically contoured distribution is proportional to that of the multivariate normal distribution. So the following likelihood ratio criteria in the multivariate elliptically contoured distribution are the same as those in the multivariate normal case:

(1) The criterion for testing for lack of correlation among the sets ($q = 1$):

$$r_1 = \frac{|\mathbf{W}|}{\prod_{j=1}^{k} |\mathbf{W}_{jj}|}. \tag{6.16}$$

(2) The criterion for testing the hypothesis that a mean vector is equal to a given vector μ_0 $(q = k = 1)$:

$$\tau_2 = \frac{|\mathbf{W}|}{|\mathbf{W}_0|} = \frac{1}{1 + T^2/(n-1)}. \tag{6.17}$$

where $\mathbf{W}_0 = \sum_{j=1}^{n} (\mathbf{x}_{(j)} - \mu_0)(\mathbf{x}_{(j)} - \mu_0)'$ and T^2 is the Hotelling T^2-statistic (see Chapter 4.)

(3) The criterion for testing the hypothesis of equality of covariance matrices $(k = 1)$:

$$\tau_3 = \prod_{i=1}^{q} \left[\left(\frac{|\mathbf{W}_i|}{\left|\sum_{j=1}^{q} \mathbf{W}_j\right|} \right)^{\frac{1}{2}n_i} \left(\frac{n}{n_i} \right)^{\frac{1}{2}pn_i} \right]. \tag{6.18}$$

(4) The criterion for testing equality of several means $(k = 1)$:

$$\tau_4 = \frac{|\mathbf{W}|}{\left|\sum_{j=1}^{q} \mathbf{W}_j\right|}. \tag{6.19}$$

(5) The criterion for testing equality of several means and covariance matrices $(k = 1)$:

$$\tau_5 = \prod_{i=1}^{q} \left(\frac{|\mathbf{W}_i|}{|\mathbf{W}|} \right)^{\frac{1}{2}n_i} \left(\frac{n}{n_i} \right)^{\frac{1}{2}pn_i}. \tag{6.20}$$

(6)The criterion for testing the hypothesis that a covariance matrices is proportional to a given matrix Σ_0 $(q = k = 1)$:

$$\tau_6 = \frac{|\Sigma_0^{-1}\mathbf{W}|}{(\mathrm{tr}\,\Sigma_0^{-1}\mathbf{W}/p)^p}. \tag{6.21}$$

6.2. Examples.

Section 6.1 gives six examples for finding LRT. In this section, three examples that do not fulfill the assumptions in Theorem 6.1 are given. From the results of these three examples, we get the idea of general theorem in section 6.3. The following three examples satisfy the assumption that $f(\cdot)$ attains its maximum at $\lambda_{\max}(g)$.

Example 1. *Tests of hypotheses concerning subvectors of μ.*

If $\mathbf{X}_{n\times p} \sim LEC_{n\times p}\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \Sigma, \phi\right)$ with $\Sigma > 0$ and $n \geq p$. We wish to test H_0 : $\mu_1 = 0$. The likelihood function has the following form :

$$L(\mu_1, \mu_2, \Sigma) = |\Sigma|^{-\frac{1}{2}n}\, g(\mathrm{tr}\,\Sigma^{-1}\mathbf{G}), \tag{6.22}$$

where $\mathbf{G} = \left(\mathbf{X} - \mathbf{1}_n\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}'\right)'\left(\mathbf{X} - \mathbf{1}_n\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}'\right)$.

Let $\Omega = \{(\mu, \Sigma) \mid \Sigma > 0\}$ and $\omega = \{(\mu, \Sigma) \mid \mu_1 = 0,\ \Sigma > 0\}$. It is clear that

$$\max_{\Omega} L(\mu_1, \mu_2, \Sigma) = \lambda_{\max}(g)^{-\frac{1}{2}np}\, g\left(\frac{p}{\lambda_{\max}(g)}\right)|\mathbf{W}|^{-\frac{1}{2}n}, \tag{6.23}$$

where $\mathbf{W}$ is obtained from $\mathbf{G}$ by substituting μ_i with $\bar{\mathbf{x}}_i$ $i = 1, 2$. Since

$$\max_{\omega} L(\mu_1, \mu_2, \Sigma) = \max_{\mu_2}\max_{\Sigma} L(0, \mu_2, \Sigma), \tag{6.24}$$

using a similar argument as the proof of Theorem 6.1, we get

$$\max_{\Sigma} L(0, \mu_2, \Sigma) = \lambda_{\max}(g)^{-\frac{1}{2}np}\, g\left(\frac{p}{\lambda_{\max}(g)}\right)|\mathbf{H}|^{-\frac{1}{2}n}, \tag{6.25}$$

where

$$\mathbf{H} = \left(\mathbf{X} - \mathbf{1}_n \begin{pmatrix} \mathbf{0} \\ \mu_2 \end{pmatrix}'\right)' \left(\mathbf{X} - \mathbf{1}_n \begin{pmatrix} \mathbf{0} \\ \mu_2 \end{pmatrix}'\right)$$

$$= \mathbf{W} + n \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_2 - \mu_2 \end{pmatrix} \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_2 - \mu_2 \end{pmatrix}'.$$

Thus

$$(\bar{\mathbf{x}}_1', (\bar{\mathbf{x}}_2 - \mu_2)') = (\bar{\mathbf{x}}_1', (\bar{\mathbf{x}}_2 - \mu_2)' - \bar{\mathbf{x}}_1'\mathbf{W}_{11}^{-1}\mathbf{W}_{12}) \begin{pmatrix} \mathbf{I} & \mathbf{W}_{11}^{-1}\mathbf{W}_{12} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{6.26}$$

$$\begin{pmatrix} \mathbf{I} & \mathbf{W}_{11}^{-1}\mathbf{W}_{12} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{W}_{11}^{-1}\mathbf{W}_{12} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{6.27}$$

and

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{W}_{21}\mathbf{W}_{11}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\mathbf{W}_{11}^{-1}\mathbf{W}_{12} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{W}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{22\cdot 1} \end{pmatrix}, \tag{6.28}$$

imply that

$$|\mathbf{H}| = \left| \mathbf{W} + n \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_2 - \mu_2 \end{pmatrix} \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_2 - \mu_2 \end{pmatrix}' \right|$$

$$= |\mathbf{W}| \left[1 + n(\bar{\mathbf{x}}_1', (\bar{\mathbf{x}}_2 - \mu_2)')\mathbf{W}^{-1} \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_2 - \mu_2 \end{pmatrix} \right]$$

$$= |\mathbf{W}| \left[1 + n(\bar{\mathbf{x}}_1', (\bar{\mathbf{x}}_2 - \mu_2)' - \bar{\mathbf{x}}_1'\mathbf{W}_{11}^{-1}\mathbf{W}_{12}) \begin{pmatrix} \mathbf{W}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{22\cdot 1}^{-1} \end{pmatrix} \right.$$

$$\times \begin{pmatrix} \bar{\mathbf{x}}_1' \\ (\bar{\mathbf{x}}_2 - \mu_2) - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1 \end{pmatrix} \Bigg]$$

$$= |\mathbf{W}| \, [1 + n\{\bar{\mathbf{x}}_1'\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1 + [(\bar{\mathbf{x}}_2 - \mu_2) - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1]'\mathbf{W}_{22\cdot 1}^{-1} \tag{6.29}$$

$$\times \, [(\bar{\mathbf{x}}_2 - \mu_2) - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1]\}] \, ,$$

where $\mathbf{W} = \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix}$ and $\mathbf{W}_{22\cdot 1} = \mathbf{W}_{22} - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\mathbf{W}_{12}$. The maximum likelihood estimate $\hat{\mu}_2$ of μ_2 is

$$\hat{\mu}_2 = \bar{\mathbf{x}}_2 - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1. \tag{6.30}$$

So

$$\max_{\omega} L(\mu_1, \mu_2, \Sigma) = \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) |\mathbf{W}|^{-\frac{1}{2}n} \left(1 + n\bar{\mathbf{x}}_1'\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1\right)^{-\frac{1}{2}n}, \tag{6.31}$$

and the likelihood ratio criterion is

$$\begin{aligned} \frac{\max_{\omega} L(\mu_1, \mu_2, \Sigma)}{\max_{\Omega} L(\mu_1, \mu_2, \Sigma)} &= \frac{|\mathbf{W}|^{\frac{1}{2}n}}{|\mathbf{W}|^{\frac{1}{2}n} \left(1 + n\bar{\mathbf{x}}_1'\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1\right)^{\frac{1}{2}n}} \\ &= \left(\frac{1}{1 + n\bar{\mathbf{x}}_1'\mathbf{W}_{11}^{-1}\bar{\mathbf{x}}_1}\right)^{\frac{1}{2}n}. \end{aligned} \tag{6.32}$$

Example 2.

If $\mathbf{X}_{n\times p} \sim LEC_{n\times p}(\mu, \Sigma, \phi)$ with $\Sigma > 0$ and characteristic roots λ_i, $i = 1, \ldots, p$, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$. We wish to test $H_0 : \lambda_{i+1} = \cdots = \lambda_{i+k}$, $i + k \leq p$. Since $\Sigma > 0$, there exists an orthogonal matrix Γ such that $\Sigma = \Gamma \Lambda \Gamma'$ where Λ is an diagonal matrix with the (ℓ, ℓ)th elements λ_ℓ, $\ell = 1, \ldots, p$ and the likelihood function has the form

$$L(\mu, \Gamma, \Lambda) = \prod_{j=1}^{p} \lambda_j^{-\frac{1}{2}n} g(\operatorname{tr} \Lambda^{-1}\mathbf{H}), \tag{6.33}$$

where $\mathbf{H} = \Gamma' \mathbf{G} \Gamma$.

Let $\Omega = \{(\mu, \Sigma) \mid \Sigma > 0\}$ and $\omega = \{(\mu, \Sigma) \mid \lambda_{i+1} = \cdots = \lambda_{i+k}\}$. It is clear that

$$\max_{\Omega} L(\mu, \Sigma) = \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) |\mathbf{W}|^{-\frac{1}{2}n}. \tag{6.34}$$

Since

$$\max_{\omega} L(\mu, \Sigma) = \max_{\substack{\mu, \Gamma \in O_p, \\ \Lambda \in \mathcal{M}_{(p,i,k)}}} L(\mu, \Gamma, \Lambda), \tag{6.35}$$

where O_p is the class of p-dimensional orthogonal matrix, and $\mathcal{M}_{(p,i,k)} = \{\Lambda \mid 0 < \lambda_p < \cdots < \lambda_{i+k+1} < \lambda_{i+k} = \cdots = \lambda i + 1 < \lambda_i < \cdots < \lambda_1\}$, by Theorem 6.1,

$$\begin{aligned}
\max_{\Lambda \in \mathcal{M}_{(p,i,k)}} L(\mu, \Gamma, \Lambda) &= \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \left[\left(\frac{1}{k}\sum_{j=1}^{k} h_{i+j\ i+j}\right)^{k} \prod_{\substack{i+k+1 \le j \le p \\ 1 \le j \le i}} h_{jj}\right]^{-\frac{1}{2}n} \\
&= \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \left[\left(\frac{1}{k}\sum_{j=1}^{k} \gamma_{i+j}' \mathbf{G} \gamma_{i+j}\right)^{k}\right. \\
&\qquad \left. \prod_{\substack{i+k+1 \le j \le p \\ 1 \le j \le i}} \gamma_j' \mathbf{G} \gamma_j \right]^{-\frac{1}{2}n},
\end{aligned} \tag{6.36}$$

where $\mathbf{H} = (h_{ij}) = (\gamma_i' \mathbf{G} \gamma_j)$, $\Gamma = (\gamma_1, \cdots, \gamma_p)$, and $h_{pp} \le h_{p-1\ p-1} \le \cdots \le h_{11}$. Therefore

$$\mathbf{G} = \mathbf{W} + n(\bar{\mathbf{x}} - \mu)(\bar{\mathbf{x}} - \mu)', \tag{6.37}$$

$$
\begin{aligned}
h_{jj} &= \gamma_j' \mathbf{G} \gamma_j \\
&= \gamma_j' (\mathbf{W} + n(\bar{\mathbf{x}} - \mu)(\bar{\mathbf{x}} - \mu)')\gamma_j \\
&= \gamma_j' \mathbf{W} \gamma_j + n\gamma_j' (\bar{\mathbf{x}} - \mu)(\bar{\mathbf{x}} - \mu)' \gamma_j \qquad (6.38) \\
&\geq \gamma_j' \mathbf{W} \gamma_j,
\end{aligned}
$$

and

$$
\max_{\mu} \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \left[\left(\frac{1}{k} \sum_{j=1}^{k} \gamma_{i+j}' \mathbf{G} \gamma_{i+j} \right)^{k} \prod_{\substack{i+k+1 \leq j \leq p \\ 1 \leq j \leq i}} \gamma_j' \mathbf{G} \gamma_j \right]^{-\frac{1}{2}n},
$$

imply that

$$
= \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \left[\left(\frac{1}{k} \sum_{j=1}^{k} \gamma_{i+j}' \mathbf{W} \gamma_{i+j} \right)^{k} \prod_{\substack{i+k+1 \leq j \leq p \\ 1 \leq j \leq i}} \gamma_j' \mathbf{W} \gamma_j \right]^{-\frac{1}{2}n}. \qquad (6.39)
$$

Let $\mathbf{F} = \Gamma' \mathbf{W} \Gamma = (\gamma_i' \mathbf{G} \gamma_j) = (f_{ij})$, $f_{pp} \leq \cdots \leq f_{11}$, and $0 \leq r_p \leq \cdots \leq r_1$ be the characteristic roots of $\mathbf{W}$. Then, by Theorem 9.B.1 due to Marshall and Olkin (1979), $\mathbf{f} \prec \mathbf{r}$ ($\mathbf{r}$ majorizes $\mathbf{f}$) on $\mathcal{D}_{\mathrm{p}} = \{ \mathbf{x} \in \Re^{\mathrm{p}} \ : \ \mathbf{x}_1 \geq \cdots \geq \mathbf{x}_{\mathrm{p}} \}$, where $\mathbf{f} = (f_{11}, \cdots, f_{pp})$ and $\mathbf{r} = (r_1, \cdots, r_p)$. Define

$$
\psi(\mathbf{x}) = \left(\frac{1}{k} \sum_{j=1}^{k} x_{i+j} \right)^{k} \prod_{\substack{i+k+1 \leq j \leq p \\ 1 \leq j \leq i}} x_j, \qquad (6.40)
$$

where $\mathbf{x} \in \Re^{p}$. Then ψ is continuous on $\mathcal{D}_{\mathrm{p}}$ and continuously differentiable on the interior of $\mathcal{D}_{\mathrm{p}}$. Since

$$\psi_\ell(\mathbf{x}) = \frac{\partial\psi(\mathbf{x})}{\partial x_\ell}$$

$$= \begin{cases} \left(\dfrac{1}{k}\displaystyle\sum_{j=1}^{k} x_{i+j}\right)^{k} \displaystyle\prod_{\substack{i+k+1\le j\le p \\ 1\le j\le i \\ j\neq\ell}} x_j & i+k+1\le\ell\le p \quad \text{or} \quad 1\le\ell\le i \\ \\ \left(\dfrac{1}{k}\displaystyle\sum_{j=1}^{k} x_{i+j}\right)^{k-1} \displaystyle\prod_{\substack{i+k+1\le j\le p \\ 1\le j\le i}} x_j & i+1\le\ell\le i+k \end{cases}, \tag{6.41}$$

is decreasing in $\ell = 1, \dots, p$, by Theorem 3.A.3 due to Marshall and Olkin (1979),

$$\begin{aligned} &\max_{f_{pp}\le\cdots\le f_{11}} \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \left[\left(\frac{1}{k}\sum_{j=1}^{k} f_{i+j\ i+j}\right)^{k} \prod_{\substack{i+k+1\le j\le p \\ 1\le j\le i}} f_{jj}\right]^{-\frac{1}{2}n} \\ &= \lambda_{\max}(g)^{-\frac{1}{2}np} g\left(\frac{p}{\lambda_{\max}(g)}\right) \left[\left(\frac{1}{k}\sum_{j=1}^{k} r_{i+j}\right)^{k} \prod_{\substack{i+k+1\le j\le p \\ 1\le j\le i}} r_j\right]^{-\frac{1}{2}n}. \end{aligned} \tag{6.42}$$

Thus the likelihood ratio criterion is

$$\begin{aligned} \frac{\max_\omega L(\mu,\Sigma)}{\max_\Omega L(\mu,\Sigma)} &= \left(\frac{\prod_{j=1}^{p} r_j}{\left(\frac{1}{k}\sum_{j=1}^{k} r_{i+j}\right)^{k} \prod_{\substack{i+k+1\le j\le p \\ 1\le j\le i}} r_j}\right)^{\frac{1}{2}n} \\ &= \left(\frac{\prod_{j=1}^{k} r_{i+j}}{\left(\frac{1}{k}\sum_{j=1}^{k} r_{i+j}\right)^{k}}\right)^{\frac{1}{2}n}. \end{aligned} \tag{6.43}$$

Example 3. *Circularly symmetric model*

If $\mathbf{X}_{n\times pk} \sim LEC_{n\times pk}(\mu, \Sigma, \phi)$ with $\Sigma > 0$ and $n \geq pk$. Consider the following four hypotheses :

$$H_1 \; : \; \Sigma = \mathrm{diag}(\Sigma_0, \cdots, \Sigma_0),$$

$$H_2 \; : \; \Sigma = \begin{pmatrix} \Sigma_0 & \Sigma_1 & \cdots & \Sigma_1 & \Sigma_1 \\ \Sigma_1 & \Sigma_0 & \cdots & \Sigma_1 & \Sigma_1 \\ \vdots & \vdots & & \vdots & \vdots \\ \Sigma_1 & \Sigma_1 & \cdots & \Sigma_0 & \Sigma_1 \\ \Sigma_1 & \Sigma_1 & \cdots & \Sigma_1 & \Sigma_0 \end{pmatrix},$$

$$H_3 \; : \; \Sigma = \begin{pmatrix} \Sigma_0 & \Sigma_1 & \cdots & \Sigma_{p-2} & \Sigma_{p-1} \\ \Sigma_{p-1} & \Sigma_0 & \cdots & \Sigma_{p-3} & \Sigma_{p-2} \\ \vdots & \vdots & & \vdots & \vdots \\ \Sigma_2 & \Sigma_3 & \cdots & \Sigma_0 & \Sigma_1 \\ \Sigma_1 & \Sigma_2 & \cdots & \Sigma_{p-1} & \Sigma_0 \end{pmatrix},$$

$$H_4 \; : \; \Sigma > 0,$$

where Σ_i, $i = 0, 1, \ldots, p-1$ are $k \times k$ symmetric and Σ is $pk \times pk$ positive definite.

If Σ is circularly symmetric, then

$$\Sigma = (\mathbf{W}_0 \otimes \Sigma_0) + (\mathbf{W}_1 \otimes \Sigma_1) + \cdots + (\mathbf{W}_{p-1} \otimes \Sigma_{p-1}), \tag{6.44}$$

where $\mathbf{W}_0 = \mathbf{I}$, $\mathbf{W}_j = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{p-j} \\ \mathbf{I}_j & \mathbf{0} \end{pmatrix}$, $j = 1, \ldots p-1$, and $\mathbf{W}_j = \mathbf{W}_1{}^j$. From the theorem in Olkin (1972), there exists $\Gamma \in O_p$, with elements r_{jk} given by

$$r_{jk} = p^{-\frac{1}{2}} \left\{ \sin\left(2\pi \frac{(j-1)(k-1)}{p}\right) + \cos\left(2\pi \frac{(j-1)(k-1)}{p}\right) \right\}, \tag{6.45}$$

such that $(\Gamma \otimes \mathbf{I})\Sigma(\Gamma' \otimes \mathbf{I}) = \mathrm{diag}(\Psi_1, \cdots, \Psi_p) \equiv \mathbf{D}_\Psi$ where Ψ_j, $j = 1, \ldots, p$, are $k \times k$ positive

definite. The parameter space for each of the hypotheses H_1–H_4 becomes

$$\omega_1 = \{\mathbf{D}_\Psi \mid \Psi_1 = \cdots = \Psi_p > 0, \text{given } \Psi_j = \Psi_{p-j+2},\ j = 2, \cdots, p\},$$

$$\omega_2 = \{\mathbf{D}_\Psi \mid \Psi_1 > 0,\ \Psi_2 = \cdots = \Psi_p > 0, \text{given } \Psi_j = \Psi_{p-j+2},\ j = 2, \cdots, p\},$$

$$\omega_3 = \{\mathbf{D}_\Psi \mid \Psi_1 > 0,\ \Psi_j = \Psi_{p-j+2} > 0,\ j = 2, \cdots, p\},$$

$$\omega_4 = \{\Sigma \mid \Sigma > 0\}.$$

Let $\mathbf{V} = (\Gamma \otimes \mathbf{I})\mathbf{S}(\Gamma' \otimes \mathbf{I}) = (\mathbf{V}_{ij}),\ i = 1, \ldots, p,\ j = 1, \ldots, p.$

When $p = 2m$,

$$(\mathbf{V}_1, \cdots, \mathbf{V}_{m+1}) = (\mathbf{V}_{11}, \mathbf{V}_{22} + \mathbf{V}_{pp}, \cdots, \mathbf{V}_{mm} + \mathbf{V}_{m+2\ m+2}, \mathbf{V}_{m+1\ m+1}), \tag{6.46}$$

while $p = 2m + 1$,

$$(\mathbf{V}_1, \cdots, \mathbf{V}_{m+1}) = (\mathbf{V}_{11}, \mathbf{V}_{22} + \mathbf{V}_{pp}, \cdots, \mathbf{V}_{m+1\ m+1} + \mathbf{V}_{m+2\ m+2}), \tag{6.47}$$

where $\mathbf{V}_j = \mathbf{V}_{p-j+2},\ j = 2, \ldots, p$, and $\mathbf{S} = (\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')'(\mathbf{X} - \mathbf{1}_n\bar{\mathbf{x}}')$.

Under ω_1, ω_2, ω_3, the likelihood function has the form

$$L(\mu, \Psi_1, \cdots, \Psi_p) = \prod_{i=1}^{p} |\Psi_i|^{-\frac{1}{2}n} g(\operatorname{tr} \mathbf{D}_\Psi^{-1}(\Gamma \otimes \mathbf{I})\mathbf{G}(\Gamma' \otimes \mathbf{I})), \tag{6.48}$$

using a similar argument as the proof of Theorem 6.1,

$$\max_{\omega_1} L(\mu, \Psi_1, \cdots, \Psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g\left(\frac{pk}{\lambda_{\max}(g)}\right) \left|\frac{1}{p}\sum_{i=1}^{p} \mathbf{V}_{ii}\right|^{-\frac{1}{2}np}, \tag{6.49}$$

$$\max_{\omega_2} L(\mu, \Psi_1, \cdots, \Psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g\left(\frac{pk}{\lambda_{\max}(g)}\right) |\mathbf{V}_{11}|^{-\frac{1}{2}n} \left|\frac{1}{p-1}\sum_{i=2}^{p} \mathbf{V}_{ii}\right|^{-\frac{1}{2}n(p-1)}, \tag{6.50}$$

$$\max_{\omega_3} L(\mu, \Psi_1, \cdots, \Psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g\left(\frac{pk}{\lambda_{\max}(g)}\right) 2^{2k(p-m-1)} \prod_{i=1}^{p} |\mathbf{V}_i|^{-\frac{1}{2}n}, \tag{6.51}$$

$$\max_{\omega_4} L(\mu, \Psi_1, \cdots, \Psi_p) = \lambda_{\max}(g)^{-\frac{1}{2}npk} g\left(\frac{pk}{\lambda_{\max}(g)}\right) |\mathbf{V}|^{-\frac{1}{2}n}, \tag{6.52}$$

and the likelihood ratio criterion is same as the one in the multivariate elliptically contoured case.

6.3. General theorem.

From three examples in section 6.2, we can derive the general theorem for finding MLE. First, let's generate the assumption for this general theorem.

Assumption B.

Let $\mathbf{x}_{N\times 1}$ have the density $|\Sigma|^{-\frac{1}{2}} g((\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu))$ for $\Sigma_{N\times N}$ positive definite and $\mu_{N\times 1}$. Let $\omega_0 \equiv \omega_m \times \omega_e$, where $\omega_m \subset \Re^N$ and ω_e is a subset of positive definite matrices such that if $\Sigma \in \omega_e$, then $\alpha\Sigma \in \omega_e$ for every $\alpha > 0$. Suppose that the MLE's under normality, $\tilde{\mu} \in \omega_m$ and $\tilde{\Sigma} \in \omega_e$, exist, are unique. $\tilde{\Sigma}$ is positive definite (with probability 1), and let the maximum of likelihood function be $\tilde{L}$.

Theorem 6.2.

Under the assumption B, if $f(\lambda) \equiv \lambda^{-\frac{1}{2}N} g\left(\frac{1}{\lambda}\right)$ atttains its maximum at some finite $\lambda_{\max}(g)$, then $\hat{\mu} = \tilde{\mu}$ and $\hat{\Sigma} = N\lambda_{\max}(g)\tilde{\Sigma}$ are maximum likelihood estimates for $g(\cdot)$ and

the corresponding maximum of likelihood function is

$$\frac{\lambda_{\max}(g)^{-\frac{1}{2}N} g\left(\frac{1}{\lambda_{\max}(g)}\right)}{N^{\frac{1}{2}N}(2\pi)^{-\frac{1}{2}N}} e^{\frac{1}{2}N} \tilde{L}. \tag{6.53}$$

Proof.

By definition, the likelihood function has the form

$$L \equiv L(\mu, \Sigma) = |\Sigma|^{-\frac{1}{2}} g(\operatorname{tr} \Sigma^{-1} \mathbf{G}), \tag{6.54}$$

where $\mathbf{G} = (\mathbf{x} - \mu)(\mathbf{x} - \mu)'$. Since $\Sigma > 0$, there exists an orthogonal matrix Γ such that $\Sigma = \Gamma\Lambda\Gamma'$, where Λ is a diagonal matrix with diagonal elements λ_i, $i = 1, \ldots, N$, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$ are the characteristic roots of Σ. Therefore

$$L = \left[\prod_{i=1}^{N} \lambda_i^{-\frac{1}{2}}\right] g\left(\sum_{i=1}^{N} \lambda_i^{-1} h_{ii}\right), \tag{6.55}$$

where $\mathbf{H} = \Gamma \mathbf{G} \Gamma' = (h_{ij})$.

Let $\frac{1}{\lambda} = \sum_{i=1}^{N} \lambda_i^{-1} h_{ii}$, $\beta_j = \lambda \frac{h_{jj}}{\lambda_j}$, $j = 1, \ldots, N-1$, then

$$L = \left[\prod_{i=1}^{N-1} \left(\frac{\beta_i}{h_{ii}}\right)^{-\frac{1}{2}}\right] \left(\frac{1 - \sum_{i=1}^{N-1} \beta_i}{h_{NN}}\right)^{-\frac{1}{2}} \lambda^{-\frac{1}{2}N} g\left(\frac{1}{\lambda}\right). \tag{6.56}$$

Define $\omega_\lambda^* = \{(\Gamma, \beta_1, \cdots, \beta_{N-1}) \ : \ \Sigma \in \omega_c\}$. Let $\Sigma^* = \alpha\Sigma$, $\alpha > 0$, then $\lambda_i^* = \alpha\lambda_i$, $i = 1, \ldots, N$ and $\lambda^* = \alpha\lambda$. Therefore the range of λ is $(0, \infty)$ and ω_λ^* does not depend on

λ (defined as ω^*). So

$$\max_{\substack{\mu\in\omega_m\\ \Sigma\in\omega_c}} L(\mu,\Sigma)$$

$$= \max_{\substack{\mu\in\omega_m\\ (\Gamma,\beta_1,\cdots,\beta_{N-1})\in\omega^*}} \left[\prod_{i=1}^{N-1}\left(\frac{\beta_i}{h_{ii}}\right)^{-\frac{1}{2}}\right]\left(\frac{1-\sum_{i=1}^{N-1}\beta_i}{h_{NN}}\right)^{-\frac{1}{2}} \max_{\lambda\in(0,\infty)} \lambda^{-\frac{1}{2}N} g\left(\frac{1}{\lambda}\right) \tag{6.57}$$

$$= \lambda_{\max}(g)^{-\frac{1}{2}N} g\left(\frac{1}{\lambda_{\max}(g)}\right) \max_{\substack{\mu\in\omega_m\\ (\Gamma,\beta_1,\cdots,\beta_{N-1})\in\omega^*}} \left[\prod_{i=1}^{N-1}\left(\frac{\beta_i}{h_{ii}}\right)^{-\frac{1}{2}}\right]\left(\frac{1-\sum_{i=1}^{N-1}\beta_i}{h_{NN}}\right)^{-\frac{1}{2}}.$$

Since $\max_{\substack{\mu\in\omega_m\\ (\Gamma,\beta_1,\cdots,\beta_{N-1})\in\omega^*}} \left[\prod_{i=1}^{N-1}\left(\frac{\beta_i}{h_{ii}}\right)^{-\frac{1}{2}}\right]\left(\frac{1-\sum_{i=1}^{N-1}\beta_i}{h_{NN}}\right)^{-\frac{1}{2}}$ does not depend on g, and under normality, $g(x) = (2\pi)^{-\frac{1}{2}N}e^{-\frac{1}{2}x}$ and $\lambda_{\max}(g) = \frac{1}{N}$, it is clearly that $\hat{\mu} = \tilde{\mu}$, $\hat{\Sigma} = N\lambda_{\max}(g)\tilde{\Sigma}$, and the corresponding maximum of likelihood function is

$$\frac{\lambda_{\max}(g)^{-\frac{1}{2}N} g\left(\frac{1}{\lambda_{\max}(g)}\right)}{N^{\frac{1}{2}N}(2\pi)^{-\frac{1}{2}N}} e^{\frac{1}{2}N}\tilde{L}. \tag{6.58}$$

∎

Combining the above theorem and Lemma 6.1, we have the following corollary.

Corollary 6.2.

Under the same assumptions as Theorem 6.2, if $g(\cdot)$ is continuous and $E(R^2) < \infty$ where $R \leftrightarrow \phi \in \Phi_N$, then $\hat{\mu} = \tilde{\mu}$, $\hat{\Sigma} = N\lambda_{\max}(g)\tilde{\Sigma}$, and the corresponding maximum of likelihood function is (6.58)

Corollary 6.3.

Let $\mu \in \Omega_m \subset \Re^N$ and $\Sigma \in \Omega_c$, satisfying the conditions of Theorem 6.2, and let the null hypothesis H be $\mu \in \omega_m \subset \Omega_m$ and $\Sigma \in \omega_c \subset \Omega_c$. Then the likelihood ratio criterion is independent of ϕ.

Proof.

The LRC is

$$\left(\frac{\hat{L}_\omega}{\hat{L}_\Omega}\right) = \left(\frac{\tilde{L}_\omega}{\tilde{L}_\Omega}\right), \tag{6.59}$$

from Theorem 6.2, which does not depend on R (i.e., ϕ.) ∎

Remarks :

(1) For every random matrix $\mathbf{X}$ *distributed as multivariate elliptically contoured distribution, we can write it in a vector form which has an elliptically contoured distribution and a density function same as the one in Theorem 6.2. Hence Theorem 6.2 is more general than Theorem 6.1. and the three examples in section 2 also follow from Theorem 6.2.*

(2) Since, from (1.6),

$$\begin{aligned} \mathrm{Cov}(\mathbf{x}) &= E(R^2\mathbf{D}'\mathbf{u}^{(N)}\mathbf{u}^{(N)'}\mathbf{D}) \\ &= E(R^2)\mathbf{D}'\,\mathrm{Cov}(\mathbf{u}^{(N)})\mathbf{D}, \end{aligned} \tag{6.60}$$

$E(R^2) < \infty$ *iff each element of* $\mathrm{Cov}(\mathbf{x})$ *is finite. So the assumptions in Corollary 6.2 — R has finite second moment and $g(\cdot)$ is continuous — are quite reasonable assumptions.*

Chapter 7

NONCENTRAL DISTRIBUTIONS OF QUADRATIC FORMS

In usual multivariate analysis, we know the distribution of $\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)}$, where $\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix} \sim N_n\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathbf{I}\right)$. Anderson and Fang (1982a) derive the distribution of $\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)}$, where $\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix} \sim EC_n\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathbf{I} \;\; ;\phi\right)$, under the condition that $\mu_1 = \mathbf{0}$. In this chapter, we derive the distribution of $\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)}$ when $\mu_1 \neq \mathbf{0}$.

If $\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix} \sim EC_n\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathbf{I} \;\; ;\phi\right)$, where $\mathbf{x}^{(1)}$ is k-dimensional vector, then $\mathbf{x} \stackrel{d}{=} \mu + R\mathbf{U}^{(n)}$, where R (≥ 0) is independent of $\mathbf{U}^{(n)} = \begin{pmatrix} \mathbf{U}_1^{(n)} \\ \mathbf{U}_2^{(n)} \end{pmatrix}$, and $\mathbf{x}^{(1)} \stackrel{d}{=} \mu_1 + R\mathbf{U}_1^{(n)}$. From Lemma 2 of Cambanis, Huang and Simons (1981), $\mathbf{U}_1^{(n)} \stackrel{d}{=} R_{kn}\mathbf{U}^{(k)}$, where R_{kn} (≥ 0) and $\mathbf{U}^{(k)}$ are independent and $R_{kn}^2 \sim \text{Beta}(\frac{k}{2}, \frac{n-k}{2})$. Thus $\mathbf{x}^{(1)} \stackrel{d}{=} \mu_1 + RR_{kn}\mathbf{U}^{(k)}$ and the distribution of $\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)}$ is

$$\begin{aligned}
\text{Prob}\{\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)} \leq x\} &= \text{Prob}\{(RR_{kn}\mathbf{U}^{(k)} + \mu_1)'(RR_{kn}\mathbf{U}^{(k)} + \mu_1) \leq x\} \\
&= \text{Prob}\{R^2R_{kn}^2 + 2RR_{kn}\mathbf{U}^{(k)\prime}\mu_1 + \|\mu_1\|^2 \leq x\} \\
&= \text{Prob}\{V + 2V^{\frac{1}{2}}\mathbf{U}^{(k)\prime}\mu_1 + \|\mu_1\|^2 \leq x\},
\end{aligned} \tag{7.1}$$

where $V = R^2R_{kn}^2$. By making an orthogonal transformation, we get $\mathbf{U}^{(k)\prime}\mu_1 \stackrel{d}{=} u_1\|\mu_1\|$, where $\mathbf{U}^{(k)} = \begin{pmatrix} u_1 \\ \mathbf{U}_2 \end{pmatrix}$. Since we know the distribution of V and u_1 (see Anderson and Fang

(1982a)),

$$\text{Prob}\{\mathbf{x}^{(1)'}\mathbf{x}^{(1)} \le x\} = \text{Prob}\{V + 2V^{\frac{1}{2}}u_1\|\mu_1\| + \|\mu_1\|^2 \le x\}$$

$$= \int_0^\infty \text{Prob}\left\{u_1 \le \frac{x - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|}\right\} g(v)\,dv$$

$$= \int_0^\infty \left[\int_{-\infty}^{b} \frac{\Gamma(\frac{k}{2})}{\Gamma(\frac{k-1}{2})\pi^{\frac{1}{2}}}(1-u^2)^{\frac{k-1}{2}-1} I_{\{|u|\le 1\}}\,du\right] g(v)\,dv, \tag{7.2}$$

where $b = \frac{x-\|\mu_1\|^2-v}{2v^{\frac{1}{2}}\|\mu_1\|}$,

$$g(v) = \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{k}{2})\Gamma(\frac{n-k}{2})} v^{\frac{1}{2}k-1} \int_{v^{\frac{1}{2}}}^\infty r^{-(n-2)}(r^2 - v)^{\frac{n-k}{2}-1}\,dF(r),$$

and $R \sim F$.

For $x \ge 0$ and $v \ge 0$,

$$\frac{x - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \le 1$$

$$\Leftrightarrow x \le (v^{\frac{1}{2}} + \|\mu_1\|)^2$$

$$\Leftrightarrow x^{\frac{1}{2}} - \|\mu_1\| \le v^{\frac{1}{2}} \tag{7.3}$$

$$\Leftrightarrow (x^{\frac{1}{2}} - \|\mu_1\|)^2 \le v \quad \text{or} \quad x \le \|\mu_1\|^2,$$

and

$$\frac{x - \|\mu_1\|^2 - v}{2v^{\frac{1}{2}}\|\mu_1\|} \geq -1$$

$$\Leftrightarrow x \geq (v^{\frac{1}{2}} - \|\mu_1\|)^2$$

$$\Leftrightarrow x^{\frac{1}{2}} + \|\mu_1\| \geq v^{\frac{1}{2}} \quad \text{and} \quad v^{\frac{1}{2}} \geq \|\mu_1\| - x^{\frac{1}{2}} \tag{7.4}$$

$$\Leftrightarrow (x^{\frac{1}{2}} + \|\mu_1\|)^2 \geq v \quad \text{and} \quad (v \geq (\|\mu_1\| - x^{\frac{1}{2}})^2 \quad \text{or} \quad x \geq \|\mu_1\|^2).$$

Let $A = \{\frac{x-\|\mu_1\|^2-v}{2v^{\frac{1}{2}}\|\mu_1\|} \leq 1\}$, $B = \{\frac{x-\|\mu_1\|^2-v}{2v^{\frac{1}{2}}\|\mu_1\|} \geq -1\}$, $A_1 = \{(x^{\frac{1}{2}} - \|\mu_1\|)^2 \leq v\}$, $B_1 = \{x \leq \|\mu_1\|^2\}$, and $C_1 = \{(x^{\frac{1}{2}} + \|\mu_1\|)^2 \geq v\}$. Then

$$A \bigcap B = A_1 \bigcap C_1.$$

We have, from (7.2),

$$\begin{aligned}
\text{Prob}\{\mathbf{x}^{(1)'}\mathbf{x}^{(1)} \leq x\} &= \int_0^\infty \left[\int_{-\infty}^{b} \frac{\Gamma(\frac{k}{2})}{\Gamma(\frac{k-1}{2})\pi^{\frac{1}{2}}}(1-u^2)^{\frac{k-1}{2}-1}(I_{A\bigcap B} + I_{A^c})\,du\right] g(v)\,dv \\
&= \int_{(x^{\frac{1}{2}}-\|\mu_1\|)^2}^{(x^{\frac{1}{2}}+\|\mu_1\|)^2} \int_{-1}^{b} \frac{\Gamma(\frac{k}{2})}{\Gamma(\frac{k-1}{2})\pi^{\frac{1}{2}}}(1-u^2)^{\frac{k-1}{2}-1}\,du\, g(v)\,dv \\
&\quad + I_{\{x\geq\|\mu_1\|^2\}} \int_0^{(x^{\frac{1}{2}}-\|\mu_1\|)^2} g(v)\,dv.
\end{aligned} \tag{7.5}$$

So, from Leibniz's rule for differentiating an integral, the density of $\mathbf{x}^{(1)'}\mathbf{x}^{(1)}$ is

$$\int_{(x^{\frac{1}{2}}-\|\mu_1\|)^2}^{(x^{\frac{1}{2}}+\|\mu_1\|)^2}(2v^{\frac{1}{2}}\|\mu_1\|)^{-1}\frac{\Gamma(\frac{k}{2})}{\Gamma(\frac{k-1}{2})\pi^{\frac{1}{2}}}(1-b^2)^{\frac{k-1}{2}-1}g(v)\,dv$$

$$=\int_{(x^{\frac{1}{2}}-\|\mu_1\|)^2}^{(x^{\frac{1}{2}}+\|\mu_1\|)^2}\frac{\Gamma(\frac{k}{2})}{\Gamma(\frac{k-1}{2})\pi^{\frac{1}{2}}}(2v^{\frac{1}{2}}\|\mu_1\|)^{-(k-2)}\{[v-(x^{\frac{1}{2}}-\|\mu_1\|)^2]$$

$$\times[(x^{\frac{1}{2}}+\|\mu_1\|)^2-v]\}^{\frac{k-3}{2}}g(v)\,dv \tag{7.6}$$

$$=\int_{(x^{\frac{1}{2}}-\|\mu_1\|)^2}^{(x^{\frac{1}{2}}+\|\mu_1\|)^2}\frac{1}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})(2\|\mu_1\|)^{k-2}2^{\frac{1}{2}k}}$$

$$\times\{[v-(x^{\frac{1}{2}}-\|\mu_1\|)^2][(x^{\frac{1}{2}}+\|\mu_1\|)^2-v]\}^{\frac{k-3}{2}}e^{-\frac{1}{2}v}h(v)\,dv,$$

where

$$h(v)=\frac{\Gamma(\frac{k}{2})2^{\frac{1}{2}k}}{v^{\frac{1}{2}k-1}e^{-\frac{1}{2}v}}g(v).$$

Let $d=x+\|\mu_1\|^2$, $c=2x^{\frac{1}{2}}\|\mu_1\|$ and $w=\frac{v-d}{c}$. The density of $\mathbf{x}^{(1)'}\mathbf{x}^{(1)}$ is

$$\frac{c^{k-2}e^{-\frac{1}{2}d}}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})(2\|\mu_1\|)^{k-2}2^{\frac{1}{2}k}}\int_{-1}^{1}(1-w^2)^{\frac{k-3}{2}}e^{-\frac{1}{2}cw}h(cw+d)\,dw$$

$$=\frac{x^{\frac{k}{2}-1}e^{-\frac{1}{2}(x+\|\mu_1\|^2)}}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})2^{\frac{1}{2}k}}I^*, \tag{7.7}$$

where $I^*=\int_{-1}^{1}(1-w^2)^{\frac{k-3}{2}}e^{-\frac{1}{2}cw}h(cw+d)\,dw$. And

$$I^*=\int_0^1(1-w^2)^{\frac{k-3}{2}}e^{-\frac{1}{2}cw}h(cw+d)\,dw+\int_0^1(1-w^2)^{\frac{k-3}{2}}e^{\frac{1}{2}cw}h(-cw+d)\,dw$$

$$=\int_0^1(1-w^2)^{\frac{k-3}{2}}\left[e^{-\frac{1}{2}cw}h(cw+d)+e^{\frac{1}{2}cw}h(-cw+d)\right]\,dw. \tag{7.8}$$

Now we make a transformation $w = s^{\frac{1}{2}}$,

$$
\begin{aligned}
I^* &= \int_0^1 (1-s)^{\frac{k-3}{2}} \left[e^{-\frac{1}{2}cs^{\frac{1}{2}}} h(cs^{\frac{1}{2}} + d) + e^{\frac{1}{2}cs^{\frac{1}{2}}} h(-cs^{\frac{1}{2}} + d) \right] \frac{1}{2} s^{-\frac{1}{2}}\, ds \\
&= \frac{1}{2} \int_0^1 (1-s)^{\frac{k-3}{2}} \left[\sum_{j=0}^{\infty} \frac{(-\frac{1}{2}cs^{\frac{1}{2}})^j}{j!} h(cs^{\frac{1}{2}} + d) + \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^j}{j!} h(-cs^{\frac{1}{2}} + d) \right] s^{-\frac{1}{2}}\, ds \\
&= \frac{1}{2} \int_0^1 (1-s)^{\frac{k-3}{2}} \left\{ \left[h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j}}{(2j)!} \right. \\
&\qquad \left. + \left[h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j+1}}{(2j+1)!} \right\} s^{-\frac{1}{2}}\, ds \\
&= \frac{1}{2}\Gamma(\frac{1}{2}) \int_0^1 (1-s)^{\frac{k-3}{2}} \left\{ \left[h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] \left[\sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j}}{2^{2j} j! \Gamma(j + \frac{1}{2})} \right] s^{-\frac{1}{2}} \right. \\
&\qquad \left. + \left[h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] \frac{c}{2} \sum_{j=0}^{\infty} \frac{(\frac{1}{2}cs^{\frac{1}{2}})^{2j}}{2^{2j+1} j! \Gamma(j + \frac{3}{2})} \right\} ds \\
&= \frac{\Gamma(\frac{1}{2}) c^{\frac{1}{2}}}{4} \int_0^1 (1-s)^{\frac{k-3}{2}} s^{-\frac{1}{4}} \left\{ \left[h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] I_{-\frac{1}{2}} \left(\frac{c}{2} s^{\frac{1}{2}} \right) \right. \\
&\qquad \left. + \left[h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] I_{\frac{1}{2}} \left(\frac{c}{2} s^{\frac{1}{2}} \right) \right\} ds \\
&= \frac{\Gamma(\frac{1}{2}) x^{\frac{1}{4}} \|\mu_1\|^{\frac{1}{2}}}{2^{\frac{3}{2}}} \int_0^1 (1-s)^{\frac{k-3}{2}} s^{-\frac{1}{4}} \left\{ \left[h(cs^{\frac{1}{2}} + d) + h(-cs^{\frac{1}{2}} + d) \right] I_{-\frac{1}{2}} \left(x^{\frac{1}{2}} \|\mu_1\| s^{\frac{1}{2}} \right) \right. \\
&\qquad \left. + \left[h(-cs^{\frac{1}{2}} + d) - h(cs^{\frac{1}{2}} + d) \right] I_{\frac{1}{2}} \left(x^{\frac{1}{2}} \|\mu_1\| s^{\frac{1}{2}} \right) \right\} ds,
\end{aligned}
\tag{7.9}
$$

where I_m is the modified Bessel function of the first kind and order m.

Hence the density of $\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)}$ is

$$\frac{x^{\frac{2k-3}{4}}e^{-\frac{1}{2}(x+\|\mu_1\|^2)}\|\mu_1\|^{\frac{1}{2}}}{\Gamma(\frac{k-1}{2})2^{\frac{k+3}{2}}}\int_0^1(1-s)^{\frac{k-3}{2}}s^{-\frac{1}{4}}\left\{\left[h(cs^{\frac{1}{2}}+d)+h(-cs^{\frac{1}{2}}+d)\right]\right.$$

$$\left.\times I_{-\frac{1}{2}}\left(x^{\frac{1}{2}}\|\mu_1\|s^{\frac{1}{2}}\right)+\left[h(-cs^{\frac{1}{2}}+d)-h(cs^{\frac{1}{2}}+d)\right]I_{\frac{1}{2}}\left(x^{\frac{1}{2}}\|\mu_1\|s^{\frac{1}{2}}\right)\right\}ds, \tag{7.10}$$

where

$$h(v)=\frac{\Gamma(\frac{k}{2})2^{\frac{1}{2}k}}{v^{\frac{1}{2}k-1}e^{-\frac{1}{2}v}}g(v),$$

$d=x+\|\mu_1\|^2$, $c=2x^{\frac{1}{2}}\|\mu_1\|$, and I_m is the modified Bessel function of the first kind and order m.

Two examples are given in the following :

Example 1. *Multivariate normal distribution.*

If $\mathbf{x}=\begin{pmatrix}\mathbf{x}^{(1)}\\ \mathbf{x}^{(2)}\end{pmatrix}\sim N_n\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\mathbf{I}\right)$ and $\mathbf{x}^{(1)}$ is k-dimensional vector. When $\mu_1=0$, the density of $\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)}$ is

$$g(v)=\frac{1}{\Gamma(\frac{k}{2})2^{\frac{1}{2}k}}v^{\frac{1}{2}k-1}e^{-\frac{1}{2}v}.$$

If $\mu_1\neq 0$, then, from (7.10) ($h(v)=1$), the density of $\mathbf{x}^{(1)\prime}\mathbf{x}^{(1)}$ is

$$2\frac{x^{\frac{2k-3}{4}}e^{-\frac{1}{2}(x+\|\mu_1\|^2)}\|\mu_1\|^{\frac{1}{2}}}{\Gamma(\frac{k-1}{2})2^{\frac{k+3}{2}}}\int_0^1(1-s)^{\frac{k-3}{2}}s^{-\frac{1}{4}}I_{-\frac{1}{2}}\left(x^{\frac{1}{2}}\|\mu_1\|s^{\frac{1}{2}}\right)ds$$

$$=\frac{1}{2}\left(\frac{x}{\|\mu_1\|^2}\right)^{\frac{1}{4}(k-2)}I_{\frac{1}{2}(k-2)}\left(x^{\frac{1}{2}}\|\mu_1\|\right)\exp\left[-\frac{1}{2}(\|\mu_1\|^2+x)\right], \tag{7.11}$$

where I_m is the modified Bessel function of the first kind and order m.

Example 2. *Multivariate t-distribution.*

From Anderson and Fang (1982a), if $\mathbf{y}_n = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$ has n-dimensional multivariate t-distribution with ℓ degree of freedom, where $\mathbf{y}_1$ is k-dimensional vector and $\mu_1 = 0$, then $\frac{1}{k}\mathbf{y}_1{}'\mathbf{y}_1$ has F-distribution with degrees of freedom k and ℓ and the density of $\mathbf{y}_1{}'\mathbf{y}_1$ is

$$g(v) = \frac{\Gamma(\frac{k+\ell}{2})}{\Gamma(\frac{k}{2})\Gamma(\frac{\ell}{2})}\ell^{-\frac{k}{2}} v^{\frac{k-2}{2}}\left(1+\frac{v}{\ell}\right)^{-\frac{k+\ell}{2}}.$$

If $\mu_1 \neq 0$, then, from (7.7), the density of $\mathbf{y}_1{}'\mathbf{y}_1$ is

$$\frac{c^{k-2}}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})(2\|\mu_1\|)^{k-2}}\int_{-1}^{1}(1-w^2)^{\frac{k-3}{2}}\frac{\Gamma(\frac{k+\ell}{2})\ell^{-\frac{k}{2}}}{\Gamma(\frac{\ell}{2})\left(1+\frac{cw+d}{\ell}\right)^{\frac{k+\ell}{2}}}\,dw$$

$$= \frac{x^{\frac{k}{2}-1}\Gamma(\frac{k+\ell}{2})\ell^{\frac{\ell}{2}}}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})\Gamma(\frac{\ell}{2})}\int_{-1}^{1}(1-w^2)^{\frac{k-3}{2}}(cw+d+\ell)^{-\frac{k+\ell}{2}}\,dw$$

$$= \frac{x^{\frac{k}{2}-1}\Gamma(\frac{k+\ell}{2})\ell^{\frac{\ell}{2}}}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})\Gamma(\frac{\ell}{2})}\int_{0}^{1}(1-s)^{\frac{k-3}{2}}\left[\left(cs^{\frac{1}{2}}+d+\ell\right)^{-\frac{k+\ell}{2}}+\left(-cs^{\frac{1}{2}}+d+\ell\right)^{-\frac{k+\ell}{2}}\right]\frac{1}{2}s^{-\frac{1}{2}}\,ds$$

$$= \frac{x^{\frac{k}{2}-1}\Gamma(\frac{k+\ell}{2})\ell^{\frac{\ell}{2}}}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})\Gamma(\frac{\ell}{2})}\sum_{j=0}^{\infty}\frac{\Gamma(\frac{k+\ell}{2}+2j)}{\Gamma(\frac{k+\ell}{2})(2j)!}(d+\ell)^{-\frac{k+\ell}{2}-2j}\int_{0}^{1}(1-s)^{\frac{k-3}{2}}\left(cs^{\frac{1}{2}}\right)^{2j}s^{-\frac{1}{2}}\,ds$$

$$= \frac{x^{\frac{k}{2}-1}\Gamma(\frac{k+\ell}{2})\ell^{\frac{\ell}{2}}}{\Gamma(\frac{k-1}{2})\Gamma(\frac{1}{2})\Gamma(\frac{\ell}{2})}\sum_{j=0}^{\infty}\frac{\Gamma(\frac{k+\ell}{2}+2j)\Gamma(\frac{1}{2})}{\Gamma(\frac{k+\ell}{2})2^{2j}j!\Gamma(j+\frac{1}{2})}(d+\ell)^{-\frac{k+\ell}{2}-2j}c^{2j}\frac{\Gamma(j+\frac{1}{2})\Gamma(\frac{k-1}{2})}{\Gamma(\frac{k}{2}+j)}$$

$$= \frac{x^{\frac{k}{2}-1}\ell^{\frac{\ell}{2}}}{\Gamma(\frac{\ell}{2})}\sum_{j=0}^{\infty}\frac{\Gamma(\frac{k+\ell}{2}+2j)x^j\|\mu_1\|^{2j}}{j!(x+\|\mu_1\|^2+\ell)^{\frac{k+\ell}{2}+2j}\Gamma(\frac{k}{2}+j)}.$$

(7.12)

References

Anderson, T.W. (1984). *An introduction to multivariate statistical analysis, 2nd edition.* Wiley, New York.

Anderson, T.W., and Fang, K.T. (1982a). Distributions of quadratic forms and Cochran's theorem for elliptically contoured distributions and their applications. *Technical Report No. 53, ONR Contract N00014-75-C-0442*, Department of Statistics, Stanford University, Stanford, California.

Anderson, T.W., and Fang, K.T. (1982b). On the theory of multivariate elliptically contoured distributions and their applications. *Technical Report No. 54, ONR Contract N00014-75-C-0442*, Department of Statistics, Stanford University, Stanford, California.

Anderson, T.W., and Fang, K.T. (1982c). Maximum likelihood estimators and likelihood ratio criteria for multivariate elliptically contoured distributions. *Technical Report No. 1, ARO Contract DAAG29-82-K-0156*, Department of Statistics, Stanford University, Stanford, California.

Cambanis, S., Huang, S., and Simons, G. (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, **11,** 368-385.

Chmielewski, M.A. (1980). Invariant scale matrix hypothesis tests under elliptically symmetry. *Journal of Multivariate Analysis*, **10,** 343-350.

Dawid, A.P. (1977). Spherical matrix distributions and a multivariate model. *Journal of the Royal Statistical Society*, **B, 39,** 254-261.

Dawid, A.P. (1978). Extendibility of spherical matrix distributions. *Journal of Multivariate Analysis*, **8,** 559-566.

Ferguson, T.S. (1967). *Mathematical Statistics : A Decision Theoretic Approach.* Academic Press, New York.

Kariya, T. (1981a). A robustness property of Hotelling's $T \to 2$-problem. *Annals of Statistics*, **9**, 210-214.

Kariya, T. (1981b). Robustness of multivariate tests. *Annals of Statistics*, **9**, 1267-1275.

Marshall, A.W., and Olkin, I. (1979). *Inequalities : theory of majorization and its applications.* Academic Press, New York.

Olkin, I. (1972). Testing and estimation for structures which are circularly symmetric in blocks. In *Multivariate statistical Inference* (D.G. Kabe and R.P. Gupta, ed.), North-Holland, pp. 183-195.

Schoenberg, I.J. (1938). Metric spaces and completely monotone functions. *Annals of Mathematics*, **39**, 811-841.

TECHNICAL REPORTS

U.S. ARMY RESEARCH OFFICE - CONTRACT DAAG29-82-K-0156

1. "Maximum Likelihood Estimators and Likelihood Ratio Criteria for Multivariate Elliptically Contoured Distributions," T. W. Anderson and Kai-Tai Fang, September 1982.

2. "A Review and Some Extensions of Takemura's Generalizations of Cochran's Theorem," George Ph.H. Styan, September 1982.

3. "Some Further Applications of Finite Difference Operators," Kai-Tai Fang, September 1982.

4. "Rank Additivity and Matrix Polynomials," George P.H. Styan and Akimichi Takemura, September 1982.

5. "The Problem of Selecting a Given Number of Representative Points in a Normal Population and a Generalized Mills' Ratio," Kai-Tai Fang and Shu-Dong He, October 1982.

6. "Tensor Analysis of ANOVA Decomposition," Akimichi Takemura, November 1982.

7. "A Statistical Approach to Zonal Polynomials," Akimichi Takemura, January 1983.

8. "Orthogonal Expansion of Quantile Function and Components of the Shapiro-Francia Statistic," Akimichi Takemura, April 1983.

9. "An Orthogonally Invariant Minimax Estimator of the Covariance Matrix of a Multivariate Normal Population," Akimichi Takemura, April 1983.

10. "Relationships Among Classes of Spherical Matrix Distributions," Kai-Tai Fang and Han-Feng Chen, April 1984.

11. "A Generalization of Autocorrelation and Partial Autocorrelation Functions Useful for Identification of ARMA(p,q) Processes," Akimichi Takemura, May 1984.

12. "Methods and Applications of Time Series Analysis Part II: Linear Stochastic Models," T. W. Anderson and N. D. Singpurwalla, October 1984.

13. "Why Do Noninvertible Estimated Moving Averages Occur?" T. W. Anderson and Akimichi Takemura, November 1984.

14. "Invariant Tests and Likelihood Ratio Tests for Multivariate Elliptically Contoured Distributions," T. W. Anderson and Huang Hsu, June 1985.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 14	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) "Invariant Tests and Likelihood Ratio Tests for Multivariate Elliptically Contoured Distributions"		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER --
7. AUTHOR(s) T. W. Anderson and Huang Hsu		8. CONTRACT OR GRANT NUMBER(s) DAAG 29-82-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics - Sequoia Hall Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P-19065-M
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE June 1985
		13. NUMBER OF PAGES 75 pp.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Invariant test, likelihood ratio test, multivariate elliptically contoured distribution, admissibility locally most powerful invariant test, maximal invariant, maximum likelihood estimation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(See reverse side for abstract.)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. Abstract

The usual assumption in multivariate hypothesis testing is that the sample consists of n independent, identically distributed Gaussian p-vectors. In this dissertation this assumption is weakened by considering a class of distributions for which the vector observation are not necessarily Gaussian or independent. This class consists of the elliptically symmetric laws with densities of the form $f(X_{n\times p}) = f(tr(X-M)'(X-M)\Sigma^{-1}$. The following hypothesis testing problems are considered: testing for equality between the mean vector and a specified vector, lack of correlations among different sets, equality of covariance matrices and mean vectors, equality between the correlation coefficient and a specified number, and MANOVA. For each of the above hypotheses, invariant tests and their properties are developed. These include the uniformly most powerful test, the locally most powerful test, admissibility, and null and non-null distributions. Further, under the assumptions that $g(\cdot)$ is continuous, each element of the covariance matrix of X is finite, and the null hypothesis is scalar-invariant, it is shown that the usual normal-theory likelihood ratio tests are exactly robust for the null case under this wider class (i.e. the likelihood ratio tests, sampling from this general class, are the same as the usual normal-theory likelihood ratio tests and their null distributions are the same.)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)